Machine Leaning methods for solar irradiance forecast blending and estimation

Javier Huertas Tato

en cumplimiento parcial de los requisitos para el grado de Doctor en

Ciencia y Tecnología Informática

Universidad Carlos III de Madrid

Directora:

Inés María Galván León

Tutor:

Ricardo Aler Mur

Noviembre 2019

Esta tesis ha sido posible gracias al apoyo que he tenido la suerte de recibir. En primer lugar, por parte de mis padres, que han sido siempre mis cimientos, con los que he podido contar toda mi vida. También deseo agradecer la presencia de aquellos amigos que han perseverado en este viaje, pese a todos los obstáculos. Por último quiero dar las gracias a mis tutores, Ricardo e Inés, que han estado siempre dispuestos a ayudarme y enseñarme. Por todo y por lo que queda, muchas gracias. Esta tesis está dedicada a todas estas increíbles personas, sin ellos nada de esto habría sido posible.

ACKNOWLEDGEMENTS

This dissertation has been developed under the project PROSOL ENE2014-56126-C2 (Towards an integrated model for solar energy forecasting) in collaboration with the research group MATRAS (University of Jaen) and funded by the Ministry of Science and Innovation (Spanish Government). All the data shown in this text has been provided by MATRAS and has been used with their permission.

PUBLISHED AND PRESENTED CONTENT

This chapter contains a list of all articles contributed within the frame of this dissertation. The following works are included in the previous work.

- Huertas-Tato, J., Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Aler-Mur, R., Galvan-Leon, I., and Pozo-Vázquez, D. (2017). Automatic Cloud-Type Classification Based On the Combined Use of a Sky Camera and a Ceilometer: Cloud class using ceilometer and camera. Journal of Geophysical Research: Atmospheres, 122(20), 11,045-11,061. https://doi.org/10/gfw59f

  - Completely included in Chapters 2 and 3 (Section 3.1). Material from this source is not marked by typography or references.

- Huertas-Tato, J., Aler, R., Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Pozo-Vázquez, D., and Galván, I. M. (2018). Predicting Global Irradiance Combining Forecasting Models Through Machine Learning. En F. J. de Cos Juez, J. R. Villar, E. A. de la Cal, Á. Herrero, H. Quintián, J. A. Sáez, and E. Corchado (Eds.), Hybrid Artificial Intelligent Systems (Vol. 10870, pp. 622-633). https://doi.org/10.1007/978-3-319-92639-1_52

  - Completely included in Chapters 2 and 4. Material from this source is not marked by typography or references.

- Rodriguez-Benitez, F., Huertas-Tato, J., Arbizu-Barrena, C., Aler-Mur, R., Pozo-Vazquez, D., and Group, M. (2018, October 7). Evaluation of a short-term solar radiation ensemble forecasting system in the Iberian Peninsula. 15, 1. Hungary. https://meetingorganizer.copernicus.org/EMS2018/EMS2018-272.pdf

  - Completely included in Chapters 2 and 4. Material from this source is not marked by typography or references.

- Huertas-Tato, Javier, Aler-Mur, R., Galvan, I., Rodriguez-Benitez, F. J., Arbizu-Barrena, C., Pozo-Vazquez, D. (In Press ,2019). A short-term solar radiation forecasting system for the Iberian Peninsula. Part 2: Model blending approaches based on machine learning. Submitted to Solar Energy.

  - Completely included in Chapters 2 and 4. Material from this source is not marked by typography or references.

- Rodriguez-Benitez, F. J., Arbizu-Barrena, C., Huertas-Tato, J., Aler-Mur, R., Galvan-Leon, I., and Pozo-Vazquez, D. (In Press, 2019). A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment. Submitted to Solar Energy.

  - Partially included in Chapters 2 and 4. Material from this source is not marked by typography or references.

RELATED RESEARCH MERITS

The following works are related to this dissertation, but have not been included in the text.

- Martín-Vázquez, R., Huertas-Tato, J., Aler, R., Galván, I. M. (2018). Studying the Effect of Measured Solar Power on Evolutionary Multi-objective Prediction Intervals. En H. Yin, D. Camacho, P. Novais, and A. J. Tallón-Ballesteros (Eds.), Intelligent Data Engineering and Automated Learning – IDEAL 2018 (Vol. 11315, pp. 155-162). https://doi.org/ 10.1007/978-3-030-03496-2_18

- Huertas, J., Rodríguez-Benítez, J., Pozo, D., Aler, R., and Galván, I. M. (2017). Genetic programming to extract features from the whole-sky camera for cloud type classification. Renewable Energy and Power Quality Journal, 1(15), 132-136. https://doi.org/10/gfw59d

- Huertas Tato, J., and Centeno Brito, M. (2018). Using Smart Persistence and Random Forests to Predict Photovoltaic Energy Production. Energies, 12(1), 100. https://doi.org/10/gf4bkw

Resumen

Las energías renovables son una importante alternativa a los combustibles fósiles ante el constante avance del cambio climático. El desarrollo de estos nuevos recursos se ha acelerado en los últimos años, especialmente en el campo de energía eólica y solar. Estas fuentes energéticas han atraído una serie de desafíos de investigación que siguen en progreso de ser resueltos, con numerosas contribuciones en la última década. La labor de estimación y predicción de energía solar es integral para el desarrollo del mercado energético, ya que permite abaratar costes instrumentales y mejorar la eficiencia de la penetración de la energía solar en la mezcla energética. La predicción de energía es fundamental en el mercado energético para estimar costes y regulaciones operativas de plantas solares, aunque la intermitencia de la energía solar hace que sea una tarea difícil. Por otro lado, la estimación de radiación solar permite reemplazar herramientas de alto coste como piranómetros y pirheliómetros; o la necesidad de expertos para detectar tipos de nube.

Para la mejora de estimación se estudian dos propuestas diferentes. En primer lugar se trata de abordar el problema de clasificación de nubes, incluyendo información de ceilómetro. Esta es una herramienta que mide altura y anchura de una nube, cuyo uso nunca ha sido aplicado en la clasificación de nubes. La siguiente propuesta es la estimación de radiación directa a partir de imágenes, usando Redes Convolucionales y múltiples perspectivas, una técnica que nunca ha sido empleada para la estimación de energía solar. Para la mejora de la predicción de energía solar se propone la integración de modelos predictivos. Esta técnica consiste en la combinación de modelos predictivos existentes para obtener una predicción final mucho más precisa que las iniciales. Aunque esta no es una aproximación nueva, su exploración ha sido insuficiente para varios modelos especializados en distintos horizontes, o para predicción a corto plazo.

Dado que las nubes producen el mayor impacto entre la radiación extraterrestre y la radiación que alcanza la superficie, las imágenes de nubes son una fuente de datos valiosa para la estimación de radiación. Para estudiar la clasificación del tipo de nube se emplea un algoritmo Random Forest entrenado con

información sobre la altura y ancho de la nube, que se combina con estadísticos obtenidos a partir de imágenes. La información del ceilómetro permite mejorar notablemente los resultados incluso cuando se incluyen ejemplos de nube difíciles para expertos. Se logra predecir 10 tipos de nube con un 71.1% de precisión frente al 50.6% obtenido sin ceilómetro. Este estudio prueba que la inclusión de información del ceilómetro tiene un impacto muy positivo en los resultados.

La estimación de radiación también se puede afrontar directamente a partir de las imágenes. Para tratar este problema se han creado varios modelos usando redes convolucionales apropiadas para el análisis de imágenes. Se proponen modelos que utilizan información proveniente de una sola cámara y otro modelo con múltiples perspectivas del cielo. Además de los canales habituales utilizados en el proceso de imágenes con redes convolucionales (RGB) se incluyen varios canales adicionales: la lejanía de los píxeles al sol y los píxeles que representan nubes. Las múltiples perspectivas y canales de información adicionales mejoran notablemente las alternativas propuestas, demostrando el aporte de la red convolucional multi-perspectiva propuesta.

Existen multitud de modelos predictivos que ofrecen predicciones con capacidades diversas a distintos horizontes de predicción. En esta tesis, se propone un modelo integrador de cuatro modelos predictivos. Usando Maquinas de Vectores de Soporte para regresión se combinan de manera lineal y no-lineal los cuatro predictores, utilizando como entradas al modelo las predicciones de los cuatro predictores. Se proponen dos aproximaciones, una por horizontes, construyendo un modelo para cada horizonte de predicción, y otra general, construyendo un modelo único para todos los horizontes. Los modelos han sido evaluados con datos procedentes de cuatro localizaciones al sur de la península ibérica.

También se propone un modelo integrador regional, capaz de aportar predicciones a nivel regional en lugar de a nivel de estación. Los resultados de integración son muy positivos tanto para radiación global como directa, mostrando mejoras absolutas hasta del 15% frente a cualquier predictor tanto en rRMSE como en rMAE. A nivel regional también se obtienen mejoras del 5% para radiación global y del 10% para radiación directa. La aproximación general es especialmente destacable, haciendo uso de un único modelo, es capaz de obtener los mejores resultados en rMAE e igualar al resto de modelos de integración en rRMSE.

Abstract

Renewable energies are the leading alternative to fossil fuels, facing the constant threat of climate change. The development of these new resources has grown in the latest years, especially in the field of solar and wind energy. These renewable power sources have gathered a series of research challenges that, to this date, are still to be solved, with many contributions to this end in the last decade. The role of estimation and forecasting of solar energy is key to the development of the solar energy market, because it cheapens instrumentation costs and improve the efficiency of solar energy market participation in the power grid. The forecast of solar energy is fundamental to estimate costs and operational regulations of a solar plant, although the intermittence of solar energy makes this a difficult task. On the other hand, the estimation of solar irradiance can replace expensive measuring devices such as pyranometers or pyrheliometers; or the need of expert supervision on meteorological stations for cloud type classification.

In order to improve estimation, two proposals are studied. The first approach to estimation is the automatic classification of clouds by including ceilometer information. This is a device capable of measuring height and thickness of a cloud, information that has never been applied to cloud classification. The next proposal is the estimation of irradiance by directly analyzing images with Convolutional Networks and multiple perspectives, a never before used technique for solar energy estimation. To improve forecasting the integration of prediction models is proposed. This technique compares and combines existing predictive models to obtain a final, more accurate, prediction. Although this is not a new approach, it has never been applied to various prediction models specialized in different horizons, or for short-term forecasting.

Given that clouds produce the greatest interference between extraterrestial and surface irradiance, whole-sky cloud images are a valuable source of data for radiation estimation. To study the cloud type classification problem a Random Forest algorithm is employed. The algorithm is trained using information from cloud height and thickness, which is combined with camera im-

age features. Including cloud height and width proves to noticeably improve accuracy even when difficult to classify cloud types are included. Results for 10-class cloud classification, including multiple clouds in a single image, show 71.12%, an improvement over the 50.6% achieved without ceilometer information. This study shows the positive impact of ceilometer information in the cloud classification problem.

Irradiance estimation can also be estimated directly from camera images. To face this problem various models have been created using convolutional neural networks, a Machine Learning technique fit for image recognition. Two approaches are proposed, a model with information from a single camera and a model with multiple sky perspectives. In addition to the common RGB colour channels used in image processing, two new channels are included: the distance from a pixel to the sun and the cloudy pixels of an image. Multiple perspectives improve noticeably all alternatives proposed, proving the contribution of the multi-view convolutional network proposed.

There are many predictive models that predict with diverse capabilities at different predictive horizons. In this thesis, this process is called forecast integration (or blending). An integration model is proposed to blend four physical models from four meteorological stations at the south of the Iberian peninsula. Using support vector regression these are combined in a linear and non-linear way using the four predictors as inputs to machine learning. Two approaches are presented: a horizon approach that builds a model for each predictive horizon, and a general approach that builds a single prediction model for all horizons.

In addition, a regional model is proposed, capable of of making predictions at a regional level instead of a station level. Results from integration are very positive compared with the baseline models for global and direct irradiance. Some absolute improvements reach 15% when comparing integration models to any predictor model when rRMSE and rMAE are evaluated on global and direct irradiance. At a regional level, there are also improvements, at an absolute 5% on global radiation over the predictor models and 10% for direct irradiance. The general approach is specially remarkable because, using a single model, it can obtain the best results on rMAE and match the results of other integration models on rRMSE.

Contents

List of Figures

List of Tables

Common Acronyms

Solar Energy

- **DNI** *Direct Normal Irradiance*
- **DHI** *Diffuse Horizontal Irradiance*
- **GHI** *Global Horizontal Irradiance*
- **NWP** *Numerical Weather Prediction model*
- **WRF** *Weather and Research Forecast*
- **CIADCast** *Cloud Index and Advection diffusion Nowcasting model*
- **TSI** *Total Sky Imager*

Artificial Inteligence

- **ML** *Machine Learning*
- **ANN** *Artificial Neural Network*
- **CNN** *Convolutional Neural Network*
- **SVM** *Support Vector Machine*
- **SVR** *Support Vector Regression*
- **RF** *Random Forest*

Chapter 1

Introduction

Global warming is a harsh consequence of the systematic overuse of earth resources. The threat of climate change is undeniable as temperatures continue to increase from pre-industrial levels. As of 2019, temperatures have risen over $0.85C^o$ with already high levels of damage to our planet [Hoegh-Guldberg et al., 2018]. Fossil fuels make up a significant percentage of the problematic greenhouse gas emissions, therefore a transition to a energy grid powered by renewable resources will greatly mitigate the impact of power generation in climate change [de Coninck et al., 2018].

Renewable energies are a clean alternative to fossil fuels, with an ever-increasing share of the energy usage all around the globe [Adib et al., 2015]. Many countries are transitioning into renewable energies for this reason, but many challenges are ahead for clean energy to achieve a high share of generation in the energy market. Two of the most prolific renewable resources are wind and solar energy, obtained from atmosphere dynamics and solar irradiance respectively. In particular solar energy faces many technological challenges yet to be solved [Kabir et al., 2018], some of which are examined within this thesis.

Given the booming industry related to solar energy, an increasing need for accurate forecasting and estimation techniques has appeared. In general and particularly in the context of this thesis, *estimation* refers to the inference of a hard-to-measure meteorological variable out of already available values at a current moment. On the other hand, *forecasting* is concerned with the prediction of meteorological variables in the future. These techniques are essential to solar power generation because accurate irradiance forecasts and estimates help make optimal management decisions in the operation of individual solar panels, solar plants, the energy market and, finally, the power grid at large [Kaur et al., 2016].

Contrary to conventional power generation, solar energy is conditioned

by weather, being highly intermittent. Estimation and forecasting of solar energy are vastly complex problems due to the interference with the atmosphere. Despite repeated efforts, irradiance estimation and forecasting models fail to accurately predict cloud motion for longer than days ahead because of the extreme complexity of the atmosphere dynamics world-wide. Solar energy in particular is strongly linked with atmospheric movements because of the existence of clouds, which make up most of the variability in solar energy, influencing greatly the power generation capabilities of solar plants. The problems associated with solar energy integration into the market have been the subject of several studies in recent years [Ela et al., 2011, Brouwer et al., 2014]. The importance of this issue with solar power generation is expected to increase as the participation of solar energy increases. One particular problem is that solar energy has a lower spatial variability than the wind energy [Santos-Alamillos et al., 2012, Santos-Alamillos et al., 2014]. This makes the regional aggregation of solar energy measures more variable and chaotic. As a consequence, the power variability generated in systems with a high solar share is higher than in the case of system with high wind energy shares [Lew et al., 2011].

Despite all the described challenges, with the advent of Machine Learning techniques, new more powerful and accurate alternatives are being developed. Over the last two decades these algorithms have been used in many areas of research, taking advantage of an abundance of data. Machine learning applications for solar energy are being continuously proposed, offering noticeable improvements over traditional physical techniques [Mellit, 2008]. However, there is a great margin of improvement for estimation and short-term irradiance forecasting [Voyant et al., 2017].

1.1 Motivation

Given the difficult challenges relative to irradiance forecasting and estimation, there is significant interest on their study in the context of solar energy. These challenges have to be addressed to improve the efficiency of solar energy penetration in the energy market, and improve the overall situation of renewable energies.

In the first place, estimation can be analysed through different angles, two of which are studied in this thesis. The first approach consists in analysing cloud types. Cloud classification is relevant to irradiance estimation because, as stated before, they are the main factor in the intermittence of solar energy. Cloud types interact with irradiance in known, predictable ways, there-

fore, knowing the cloud type gives relevant information of solar irradiance. However, there is a major disadvantage with this method of irradiance estimation, which is the requirement of experts to label clouds. Experts are costly, slow and inconsistent at classifying sky images [Hoyt, 1978], disadvantages that could be solved through the use of an automatic cloud classification model. Despite all the literature about this subject, ceilometer information has never been used for the purpose of cloud classification. A ceilometer measures height and thickness of cloud layers, which are extremely relevant features to cloud type. Therefore, there is a clear need for an exploration of the ceilometer capabilities on cloud classification.

The second way explored to estimate irradiance is the analysis of sky images. Gathering solar irradiance data is difficult due to the high costs of solar instruments, such as pyranometers or pyrheliometers. This condition makes solar energy forecasting (or other challenges) inaccessible for most researchers [Coppolino, 1994]. On the other hand, sky cameras are reliable and affordable, which can be used for irradiance estimation.

There are several works in the literature addressing irradiance estimation through the use of cameras and machine learning, however some issues arise within this subject that have not yet been examined. First, the raw camera images are never used directly in combination with machine learning, no attempts have been made to estimate irradiance this way. The usual approach to irradiance estimation with camera images is to calculate human defined features. The second issue is that multiple camera perspectives are never used. Both aspects can be studied to advance the field of solar irradiance estimation.

As it has been previously mentioned, irradiance forecasting can be used to optimize the market integration of solar resources into the energy grid. There is an extensive body of work on forecasting techniques through machine learning. However, a fairly unexplored route of research is the blending of irradiance forecasts. Blending is the process by which a set of prediction models are combined with the intent to provide more accurate predictions than those made by any of the original models. The blending approach has not received proper attention in the literature, the study of short-term irradiance blending and with multiple heterogeneous models has been neglected. A need to study blending at a local scale is found. Finally, for operators, regional irradiance forecasts [Renné, 2014, Pierro et al., 2017] are more useful than local predictions. Blending techniques have barely been studied for regional aggregation, therefore new blending approaches for regions are needed.

1.2 Objectives

This thesis has an overarching goal to advance in the fields of solar energy estimation and forecasting through machine learning methods. From this main research challenge, a set of specific objectives are defined.

The *first* objective is to study the influence of ceilometer in cloud classification problems by means of machine learning. An accurate cloud classification model is to be developed. To ensure the necessity of the ceilometer information, the new information needs to discern difficult cloud types. The *second* objective is to build a Convolutional Neural Network (CNN) able to directly estimate solar irradiance out of camera images. Also, the CNN has to support simultaneous image inputs to include multiple perspectives from the same sky. The study of the capabilities of domain specific information of sky conditions in made by including additional information channels to the conventional RGB images. Both objectives belong to the main objective of improving current irradiance estimation techniques through machine learning.

The *third* objective is to build an irradiance forecast integration (or blending) model capable of forecasting irradiance at short-term horizons (15 to 360 minutes). The purpose is to automatically detect synergies in diverse forecasting models and combining their respective strengths to obtain a final, more accurate, forecast. The prediction horizon is very relevant, as there are diverse forecast models that excel at different forecasting horizons. Therefore, it is also important to study the influence of the horizon for integration models. Furthermore, to solve the need of regional irradiance forecasting, another integration model has to be built by using information from multiple locations. These integration models are meant to advance the objective of improving current irradiance forecasting techniques.

Structure

Chapter 2 reviews the literature supporting the principles explained in the introduction and examines relevant works to the contributions of this dissertation. Chapter 3 is focused on estimation, containing the cloud classification and irradiance estimation contributions, describing the proposal, data, methodology and subsequent evaluation. Chapter 4 introduces the blending models to forecast irradiance at different forecasting horizons. In here, the proposal, data, methodology and evaluation of the forecast integration contribution are explain. Both the station and regional models are present in this section. Finally, in the last chapter, the main contributions of this thesis are

also presented. Chapter 5 draws the conclusions and further lines of research.

Chapter 2

State of the art

2.1 Terminology

Solar Irradiance is the electromagnetic radiation that reaches Earth from the Sun, measured in energy per unit area (W/m^2). There are several types of irradiance and, for the prediction of photovoltaic energy, only some types are relevant to the Earth's surface. The terms described here represent the most important irradiance types for surface forecasting and estimation.

- *Direct Normal Irradiance* (DNI) represents the amount of energy at a surface perpendicular to the Sun, excluding the irradiance diffused by atmospheric conditions (aerosols, clouds, humidity...). It can be expressed as the extraterrestrial irradiance minus the diffused irradiance. This component is highly variable and difficult to predict in nature due to the complex behaviour of the atmosphere.

- *Diffuse Horizontal Irradiance* (DHI) is the irradiance diffused by the atmosphere and light rays are scattered into diverging points in the surface. Without an atmosphere to diffuse irradiance, there would be no need to measure DHI, only DNI.

- *Global Horizontal Irradiance* (GHI) is the total irradiance that reaches the Earth on a horizontal surface, including direct and diffuse irradiance. Therefore, it can be expressed as the combination of both DNI and DHI.

In addition to these measured values, the DNI, DHI and GHI have an associated theoretical *clear sky irradiance* (or extraterrestial irradiance). This value is the theoretical amount of solar irradiance that there would be if there were no interferences between the sun and the earth. Each clear sky irradiance can be expressed as a function of latitude, longitude and time.

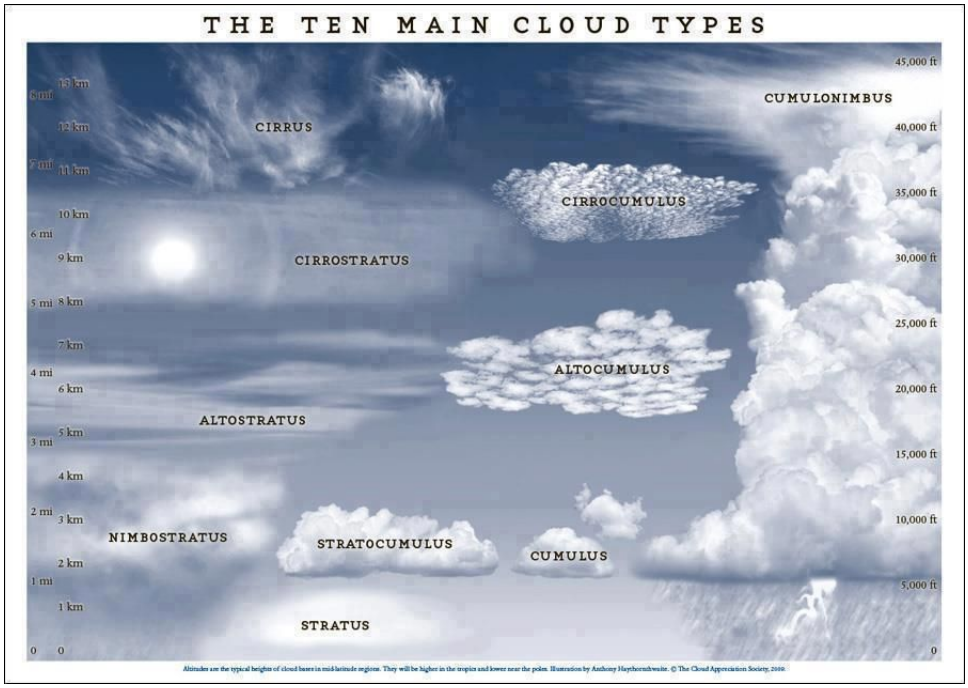

Figure 2.1: Cloud type classification by [Cloud Appreciation Society, 2011].

Given that GHI is the composition of DNI and DHI, it is not necessary to know all three measures to estimate Solar Irradiance over a horizontal surface. Only GHI and DNI are necessary to estimate irradiance for photovoltaic energy purposes. However, GHI is influenced by the DHI component of irradiance. The behaviour of DHI is tightly linked to sky conditions such as clouds.

Cloudiness plays a major role in determining GHI, and the shape, height, and texture of clouds are determining factors on how much DHI irradiance reaches Earth and how much DNI is diffused through the atmosphere. The cloud taxonomy that follows describes the standard nomenclature of the most frequent cloud types, including a brief description of each class. A visual representation of cloud types is included in Figure 2.1.

- Cumulonimbus: A cloud that spans all height levels of the lower atmosphere. They usually carry rainfall and are common in thunderstorms.

- Cirrocumulus: Very short-lived high level clouds formed mostly by ice crystals. Their shape is that of small patches of clouds and can be referred as ”cloudlets”.

- Cirrostratus: Thin clouds that form a veil at high points in the atmosphere. Their interference with the sun can produce halos if the cloud layer is thin enough.

- Cirrus: High level clouds that usually form thin strings. Can range from grey to white and can be extremely long.
- Altocumulus: Formed at mid height altitude, this type of cloud resembles the Cirrocumulus cloud. The Cloudlets are bigger and more opaque.
- Altostratus: Thick clouds that form a sheet pattern at a middle heights. They can be continuous or have small patches of clear sky.
- Nimbostratus: Amorphous and dark type of cloud, tightly related to rainfall. As the Cumulonimbus it can span multiple height levels usually from low to middle height.
- Cumulus: Flat and cotton-like clouds that appear at low levels of altitude. They are usually the base for bigger higher clouds such as Cumulonimbus.
- Stratus: This kind of clouds appears as if there was mist above surface level, at a low level of altitude. As their counterparts, the Altostratus and Cirrostratus, this kind of cloud forms a continuous layer of cloudiness.
- Stratocumulus: Large masses of round dark clouds, appearing at low level of altitude.

2.2 Solar Energy Estimation

The field of renewable energies requires many tools to be approachable. For instance, there is a need in meteorological stations for human experts to measure some specific sky features and conditions. Cloud types are not a property of the sky that can be measured by any existing tool and typically requires human expertise. Other meteorological features require specialized instruments that are frequently expensive and difficult to maintain, such as pyranometers and pyrheliometers for solar irradiance measurements. Estimation through Machine Learning techniques is a viable solution to this problem. By using inexpensive commercial hardware combined with state of the art machine learning techniques, features like irradiance, temperature, cloud type, etc, can be estimated (instead of being directly measured).

There is interest in estimating solar irradiance and cloud types for forecasting, and many estimation tasks are ongoing research challenges. Total-sky imagers (TSI) are versatile and inexpensive devices, which have been repeatedly used in estimation of meteorological and atmospheric variables [Rangarajan et al., 1984, Kreuter et al., 2009, Gauchet et al., 2012]. These devices are often

commercial cameras with slight modifications such as fish eyed lenses that are cheap to acquire. They can take pictures of the sky at any given moment and, through software computations, can be turned into estimates of irradiance [Dev et al., 2016], cloud coverage [Boers et al., 2010] or cloud type [Heinle et al., 2010].

Thus, TSIs can be used, for example, to replace pyranometers and pyrheliometers. These tools detect light wavelength fluctuations (W/m^2) to measure solar irradiance. Both GHI and DNI can be obtained from the data produced by these instruments. However their cost is often very high, disabling its extended use by researchers and solar energy power plants alike. Total sky imagers could play the same role by using available meteorological information (measured by cheap devices) and sky images. For this to be possible, a model has to be developed so that an image is transformed into an estimate. Another avenue for TSIs is complementing the information provided by other instruments. For instance, a ceilometer is a tool able to measure the cloud height and depth through LIDAR detection and can work together with TSIs for cloud type classification. Lidar devices measure distance by using intermittent pulsating lasers, measuring how much time passes between pulses the distance from the device to an object can be measured.

An issue to bear in mind when dealing with estimations is that, given they are not direct measurements, they are bound to present inaccuracies. Thus, estimations can propagate errors if they are used as real measures. The use of estimates in machine learning requires careful consideration as many algorithms are sensible to outliers and faulty data.

Next, cloud classification literature is reviewed, followed by a review of irradiance estimation study. Both problems have traditionally been approached by meteorological models, but recently machine learning approaches have repeatedly proven their accuracy.

2.2.1 Cloud classification

Scientific interest in retrieving cloud information dates many decades back and was mainly related to civil and military aviation. The attention to cloud information has grown in the framework of climate studies, since clouds play a key role in Earth's energy balance [Li et al., 2014, Wild et al., 2013]. More recently, in the field of weather forecasting, the improvement in cloud identification has emerged as a significant research field. Mainly, because clouds are involved in multiple and strong interactions within the atmosphere, and by extension with components such as irradiance, wind or temperature. Their

misrepresentation may have large impacts and implications in the accuracy of the simulations of atmospheric dynamics and therefore, of the numerical weather prediction models [Haiden et al., 2015, Pincus et al., 2011]. Lastly, the growing penetration of solar energy around the world has fostered a great interest in cloud information, since clouds are the main source of variability of the solar energy [Martínez-Chico et al., 2011, Mateos et al., 2014, Tzoumanikas et al., 2016]. In all these fields of science, the establishment of proper, accurate, and cheap cloud monitoring systems is crucial [World Meteorological Organization , 2012].

Nevertheless, the type of cloud information needed, such as cloud parameters, temporal and spatial resolution, among others, greatly varies depending on the application. In some of the above mentioned applications, information about the type of cloud is crucial. Human-reported classification is the first available continuous source of knowledge on cloud type. But the high associated cost, slow classification, low accuracy and sometimes inconsistent labelling make this source of information difficult to obtain in many countries.

Many meteorological methods have been developed for cloud classification tasks. In this regard, the use of satellite retrieval for cloud classification is a promising tool because of their spatial coverage. For instance, the "cloud-type" product of EUMETSAT Application Facility on Climate Monitoring or the APOLLO project [Kriebel et al., 2003, Wey and Schroedter-Homscheidt, 2013] reports operationally a coarse cloud classification. Nevertheless, the performance of these operational cloud-type monitoring systems is still underwhelming due to limitations of the satellite platforms.

Alternatively ground-based sky cameras systems can be used for cloud classification via feature extraction and analysis. These systems, which date one decade back, are now considered the reference for cloud cover estimates [Boers et al., 2010, Cazorla et al., 2008, Long et al., 2006]. More recently, the automatic recognition of cloud types has emerged as a possible product of these instruments. There are two basic steps for automatic cloud classification. First, the extraction of appropriate and distinctive information of the different sky conditions and cloud types from the camera images. To this end, features can be computed on the colour data from the camera channels. Particularly, these features account for characteristics such as cloud shape, texture, or cloud colour. Second, once a set of distinctive features are obtained, the cloud type assignment task relies on the use of automatic classification algorithms. These are trained and tested with human-supervised cloud-type databases. The type and number of features have increased enormously in the last years, benefiting from other fields of research, such as automatic pattern

recognition. For instance, [Calbó and Sabburg, 2008] used texture properties and the Fourier transform of the camera visible channels to classify up to eight classes of sky conditions. The methodology achieved an accuracy of about 62%.

Machine Learning Approaches

Regarding classification machine learning algorithms, the literature contains proposals ranging from artificial neural networks (ANN), to k-nearest neighbor (KNN) and support vector machines (SVM). ANNs are commonly used in the context of cloud classification. It is actually a nonlinear regression technique that can be used for classification by setting a threshold on the output(s). Typically, standard architectures with three layers are used (input/hidden/output) and in a multi-class classification context, like cloud classification, there are as many output neurons as classes. An early approach is proposed by [Lee et al., 1990] where via a single channel of the LANDSAT MSS image is processed through a Multi Layer Perceptron (MLP) to achieve 84% accuracy on 3-type classification. More recent approaches include additional feature extraction algorithms as presented by [Singh and Glennen, 2005]. Neural networks and KNNs are used to classify cloud types using different groups of features to achieve varied results on 5-type classification. [Kliangsuwan and Heednacram, 2015] proposes a new methodology, based on the fast Fourier transform, for feature extraction for cloud classification. These features are used by a MLP that achieves an overall accuracy of 90% for the automatic classification of seven clouds types.

Another common approach to cloud classification is using KNNs. KNN does not need to fit a model to the data; rather, it stores all data and classifies new instances by looking for the closest stored data instance(s). KNN does not require any adaptation for multiclass problems. The basic version of KNN may suffer more than other methods when there are many features, or some of them are irrelevant. It is also very slow for real use if the data set is large. However, there are methods for KNN that can improve both accuracy (like [Weinberger and Saul, 2009]) and speed (like kd-trees, [Wess et al., 1994]). The use of KNN is included in the proposal by [Heinle et al., 2010] where a mixed set of textural and colour features are used for the classification of up to seven cloud types, with a success rate of about 75%. In the work of [Kazantzidis et al., 2012] a multicolour criterion on sky images is developed and presented to the KNN, and results in an average performance of about 87% using seven cloud categories. [Rumi et al., 2013] includes features from the infra-red channels of a camera to the KNN, obtaining an accuracy of 90%

in the estimation of towering cumulus and cumulonimbus cloud types. A cloud classification method based on division of the image in different blocks is proposed by [Cheng and Yu, 2015]. This way the KNN was able to account for mixed clouds types in one image, obtaining an improved classification accuracy. [Wacker et al., 2015] used, as ancillary information for cloud classification, the measured longwave irradiance as an additional feature to the KNN. An improvement of up to 10% is observed, compared to the use of just the sky camera information while the average accuracy ranged from 80 to 90%.

SVMs are rising in their use within the cloud classification problem. SVMs aim to maximize the generalization capabilities by finding separation boundaries between classes that maximize the margin. They have fewer local minima issues compared to ANNs, because SVMs solve a constrained convex optimization problem, with a single global optimum. On the other hand, the most common approach to SVM trains binary classifiers. SVMs require to train as many models as classes (one-versus-rest approach) or as many as pairs of classes (one-versus-one approach), although there are also approaches that deal with multiple classes directly [Crammer and Singer, 2001]. [Taravat et al., 2015] examines both MLPs and SVM via presenting the raw image channels to each algorithm to achieve 95.07% and 93.66% accuracy respectively on a 4-class task. The work of [Zhen et al., 2015] uses cloud classification aimed at PV forecasting. A mixed set of 12 spectral and 12 textural features and runs an SVM to accurately classify 96.44% instances of a 4-class problem. [Li et al., 2016a] used a novel approach for cloud-type recognition, based on the analysis of image as a collection of patches, rather than a collection of pixels. The proposed SVM showed an accuracy of 90% for five classes of sky conditions.

Random forests (RFs) [Breiman, 2001] are a powerful machine learning technique that has seldom been used for cloud classification but is known to be among the best performers in classification tasks according to some empirical studies [Caruana et al., 2008, Caruana and Niculescu-Mizil, 2006]. In a recent work [Cheng and Lin, 2017], RFs have been used together with other algorithms (such as SVM and a Bayesian classifier) to develop a voting scheme for classifying each pixel in the image as cloud or non-cloud.

On recent years Convolutional Neural Networks (CNN) [LeCun et al., 2015] have risen in usage due to their power on image processing problems. Therefore, TSI images can be processed through CNNs to obtain predictions. In cloud classification, [Ye et al., 2017] processes cloud TSI images through a convolutional network to classify cloud types. Filtered image patterns are taken from both first and last convolutional layers to generate a vector of rep-

resentative features of the cloud. The vector is used as inputs to a SVM. This method shows very positive results on the 9 classes of the dataset presented achieving over 80% of classification accuracy.

Aside from the sky cameras, the use of ceilometers for cloud property retrieval has emerged in the last few years [Illingworth et al., 2007]. Ceilometers are single-wavelength low-powered lidars (light detection and ranging), which can provide high-frequency observations of cloud profiles, including parameters such as the cloud base height (CBH) and cloud top height (CPD e.g., cloud thickness). Unlike satellite imagery, which generally provides low-reliability CBH estimates at relatively low temporal resolution (very few samples per hour), ceilometers are able to provide an accurate description of the location of the cloud vertical boundaries with even several samples per minute [Arbizu-Barrena et al., 2015, Costa-Surós et al., 2014, Martucci et al., 2010, Viúdez-Mora et al., 2015].

2.2.2 Irradiance Estimation

Research on the estimation of solar irradiance has had historical interest over the years. Data from solar irradiance is difficult and expensive to extract due to the associated costs of pyranometers and related devices, therefore it is difficult to gather a database with solar irradiance information [Coppolino, 1994]. Alternatives to traditional instruments have been developed over the years with the help of TSIs, weather variables or Numerical Weather Predictors (NWP) to estimate irradiance directly [Dev et al., 2016, Chow et al., 2011, Alonso-Montesinos and Batlles, 2015, Alonso-Montesinos et al., 2015b] or indirectly through cloud coverage [Kreuter et al., 2009, Pfister et al., 2003]. Estimation is usually done on daily or monthly measures of energy [Bakirci, 2015, Fan et al., 2018a], however this kind of estimation does not offer enough information for the hourly energy market.

Another approach to irradiance estimation is using satellite regional imagery [Şenkal and Kuleli, 2009, Alonso-Montesinos et al., 2015a]. These estimates also use several statistical dimensionality reduction techniques to make satellite images approachable. If more granularity in time scale and locality is required TSIs can take instant images of the sky and can be more suitable for estimation. There is a need for more work in the study of smaller time intervals [Zhang et al., 2017].

As with cloud classification, irradiance has been estimated through meteorological methods, and recently through machine learning. These algorithms can be applied over simplified statistical measurements from TSIs and mete-

orological data, or to images directly through image processing techniques.

The earliest approaches to irradiance estimation appeared because of the lack of irradiance data in research centers. Lack of data drove these models to find regression expressions that approximated solar irradiance. Some early examples used available meteorological data as proposed by [Sabbagh et al., 1977], where the authors empirically find an equation. Others like [Rangarajan et al., 1984] used the cloud-cover approach to solar irradiance prediction to form a cloud fraction solar model (i.e. to estimate irradiance from the fraction of cloud-covered sky). One of the first approaches using sky cameras is presented by [Sabburg and Wong, 1999] where using processed TSI images, estimate cloud cover and correlate solar irradiance. Their method shows weaknesses when cloud-cover is badly estimated, like cirrus-type dark clouds.

Information from images is commonly used for estimation. In [Pfister et al., 2003] the authors use a sky imager, historical irradiance measurements and clear-sky irradiance. A cloud-cover estimation is retrieved from the images and all chosen parameters are fitted through least squares. Using the found expression, the inverse correlation between irradiance and cloud-cover is shown and a relation is found. The highest improvements in accuracy are made when there is low cloud-coverage. The work of [Kreuter et al., 2009] extends its research of cloud coverage into UV irradiance and shows that estimating cloud-coverage is an useful approach to irradiance estimation. Images are not uniform, and preprocessing images requires sky segmentation as is proposed by [Gauchet et al., 2012]. Applying equations with differing parameters to separate portions of the sky reports good performance on GHI and DNI estimation. TSIs are not the only kind of sky images available, and satellite imagery can be used as presented by [Alonso-Montesinos et al., 2015a]. Here, features are taken from satellite images to estimate irradiance.

Estimating solar irradiance with aerosol-covered skies, thin clouds and other sky conditions that let through some solar irradiance can be a difficult task. An opacity detection method is proposed by [Ghonima et al., 2012] which shows good performance detecting cloud thickness, however it is not tested for solar irradiance, just for cloud coverage estimation. Other sky recognition approach is presented by [Kreuter and Blumthaler, 2013]. Images are polarized and a model is proposed for studying the feasibility aerosol detection. The relation of clouds and coverage is yet to be completely refined as pointed by [Tzoumanikas and Kazantzidis, 2016]. In [Alonso-Montesinos and Batlles, 2015] the authors take advantage of the differences between sky conditions (clear and overcast skies) to improve upon the quality of estima-

tions.

Machine Learning Approaches

Machine learning methods have slowly been adopted for solar irradiance estimation. Using meteorological variables as inputs to machine learning is a common approach. For example, [Şenkal and Kuleli, 2009] uses geographical information (latitude, longitude and altitude) paired with temporal variables and the average of historical of DHI and DNI. These features are run through an ANN to estimate monthly average irradiance on Turkey. [Chu et al., 2013] proposes using an ANN with features extracted from images, and optimizing said network using a Genetic Algorithm (GA) for DNI hourly estimations. Other less common approaches may be competitive in estimation as proposed by [Salcedo-Sanz et al., 2014]. This approach uses a combination of an evolutionary algorithm called Coral Reef Optimization (CRO) and an Extreme Learning Machine (ELM) to estimate global daily solar irradiance.

Despite scarce application in irradiance estimation, some related approaches have appeared over recent years using CNNs. In their work [Siddiqui et al., 2019] proposes to combine video feeds and weather information into a multiphase prediction model. First irradiance is estimated through a convolutional neural network. The estimated irradiance is then processed through a network, a type of Recurrent Neural Network (RNN) with good memory capabilities. Another LSTM module with auxiliary information is combined with the processed irradiance and a forecast is given. Their results overcome ECMWF, GFS and using meteorological information only. Although the cited source develops an estimation technique for images, there is no direct analysis of the estimation capabilities of the proposed CNN.

2.3 Solar Energy Forecasting

In recent years the need for renewable energy has drastically increased. The energy industry has been constantly developing solar stations at an increasing rate and the management of the solar resource has become a necessity. Solar Energy Forecasting plays an integral role in the renewable energy market. However solar and wind resources are particularly hard to forecast due to the complexity of atmosphere dynamics, acting almost randomly in practice. Therefore it is considered an unsolved research problem with plenty of potential for improvement.

A wide variety of techniques can be applied to effectively and accurately forecast solar energy. Three of them are identified as follows: physical fore-

casting, time series forecasting and forecast integration. The first approach, meteorological models have been the standard to atmospheric phenomena forecast tasks for decades. Mathematical models based on meteorological data are developed to predict irradiance or any other weather property. NWPs are meteorological models that simulate weather conditions to extrapolate the desired variables. Physical models within this category can range from simple (like Smart Persistence) or hour long complex simulations (like WRF-Solar) [Jimenez et al., 2015].

More recently time series forecasting have trended in use with the aid of machine learning algorithms. This procedure takes historical data and, out of past trends, creates a regression mathematical model that, given a set of inputs, can predict irradiance. Machine learning approaches have proven repeatedly to be very effective in the solar energy forecasting problem [Mellit, 2008, Voyant et al., 2017, Zamo et al., 2014a, Zamo et al., 2014b]. These techniques are very reliant on the quality and amount of data available, as well as the chosen inputs to train the time series model.

The last identified approach is forecast tntegration (or blending), which takes a set of forecasting models predictions and combines and corrects their respective capabilities to achieve greater forecasting accuracy. This process can be assisted by machine learning techniques. For meteorology, these forecasts could be the output of either a NWP or a time series forecasting model. Blending is a complex technique to apply due to the need of already-existing forecasting models able to output forecasts which require significant time and resources to calculate.

An study of the various NWP models is presented, followed by a review of time series forecasting and forecast integration techniques in the field of renewable energies.

2.3.1 Forecasting by Physical modelling

In the development of solar energy forecasting models many prediction approaches have been proposed based on different techniques. A reference proposal is based on satellite imagery processing [Lorenz and Heinemann, 2012]. Satellite algorithms use cloud index maps to estimate the movement of clouds through Cloud Motion Vectors (CMVs), estimating the atmospheric flow and, finally, providing a final irradiance forecast. This method was first proposed by [Beyer et al., 1994] and it has been improved for years over many works continuing the use of CMVs to this day [Miller et al., 2018].

NWP models are a another short-term approach to forecasting. These

make use of mathematical representations of the oceans and atmosphere to predict any weather condition. They have been repeatedly used for solar energy forecasting [Mathiesen and Kleissl, 2011, Lara-Fanego et al., 2012, Lorenz et al., 2014] with variable effectiveness. Despite the amount of resources required in the simulations, the NWP models lack the ability to accurately reproduce cloud amount and location [Ruiz-Arias et al., 2016].

Overall comparison of NWP and Satellite models show that satellite approaches are better for the first hours, while NWP models start overcoming satellite results on farther prediction horizons. The point at where NWP overtakes Satellite differs from one study to another, some put this point at 3 hours [Wolff et al., 2016, Kühnert et al., 2013] while [Perez et al., 2010] finds this point to be around 4 hours. Some found this point at 6 hour horizon [Lorenz and Heinemann, 2012]. It is also known that satellite is unsuitable to complex terrain [Guillot et al., 2012, Arbizu-Barrena et al., 2017]. The difference exists over regions, models or time period.

2.3.2 Forecasting by means of Time Series

Time series techniques take historical data out of present and past examples to predict future values. This methodology has seen use in many fields of research including renewable energies. It has been studied exhaustively for decades and it has proven to be effective on many problems repeatedly. [De Gooijer and Hyndman, 2006]

The foundation of this kind of forecasting is the prediction of univariate time series. This technique consists in, given p past values of a variable V that changes over time, predict a new future value at a given time point $t + h$, where t represents the present time instant and h the prediction horizon. The value h represents for which moment in time the prediction made, and stands for an increment of t. This can be represented as Equation 2.1.

$$V(t+h) = f(V(t-1), V(t-2), ..., V(t-p)) \tag{2.1}$$

Where V is the variable that is to be predicted and the only variable used. The value V may stand for any type of value depending on the problem, and it can represent irradiance, economic growth, precipitations or wind speed. The input to V is always a measure of time, like t representing the current time instant, $t + h$ which is the target time instant or $t - p$ that indicates the oldest value taken. The window p represents how old can the data be and it is a decrement of t. In the context of forecasting a function f learned from the inputs, from $V(t)$ to $V(t - p)$.

However, Equation 2.1 is unable to use more than a single variable V to predict in the time instant $t + h$. For more complex problems, univariate time series forecasting is not flexible enough to achieve acceptable predictions. Multivariate time series forecasting extends this approach to include further variables as inputs. This is represented in Equation 2.2.

$$\begin{aligned} P(t+h) = f(&V_1(t-1), V_1(t-2), ..., V_1(t-p), \\ &V_2(t-1), V_2(t-2), ..., V_2(t-p), \\ &..., \\ &V_n(t-1), V_n(t-2), ..., V_n(t-p)) \end{aligned} \tag{2.2}$$

Where $V^p(t + h)$ represents the value to be predicted in the time instant $t + h$. There are n different V_k time series variables for this expression.

Any machine learning regression algorithm can be used. The role of machine learning is finding the function f that fits best the equality for multivariate time series, minimizing an error metric between real and predicted values.

Applications related to time series forecasting applied to solar energy are described in the following subsection. Wind energy forecasting is relevant due to their relation to the atmosphere, therefore an analysis of related time series forecasting applications to wind energy is included.

Time series for Solar Energy Forecasting

Forecasting techniques have been studied for years in another areas of research, however the application of forecasting to solar irradiance and energy is recent and is still developing. In their review [Voyant et al., 2017] shows that there are machine learning algorithms that are more popular than others. ANNs are frequently used, Support Vector Regression (SVR) is starting to see some use and other algorithms like RF, Extreme Gradient Boosting (XGB) or ELMs are rarely, if ever, used.

ANNs have seen repeated use in forecasting the solar resource [Yadav and Chandel, 2014]. One of the main problems of ANNs is their susceptibility to over-fitting, however as [Kashyap et al., 2015] shows, once the typical over-fitting of ANNs has been overcome the predictions are competitive.

The usual approach to forecasting solar energy with ANNs is to use meteorological and irradiance features, in a similar way to how NWPs operate. The inclusion of atmospheric data has been repeatedly applied with variable degrees of success. For example [Orjuela-Cañón et al., 2017] employs ANNs

through this method and compares multiple machine learning algorithms to show that ANNs can offer accurate predictions on short-term horizons. Both [Koca et al., 2011] and [Ozgoren et al., 2012] apply ANNs over the region of Turkey to predict solar energy. In their work [Sharma et al., 2016] presents a Mixed Wavelet Neural Network (WNN), a variant of an ANN over the region of Singapore. The prediction on mountainous areas differs from the usual forecasting, as irradiance prediction is terrain-dependant, however ANNs are able to obtain acceptable results as [Bosch et al., 2008] proposes.

Another way to forecast solar irradiance consists in including the sky state into the prediction through visual resources such as images or video feeds. The sky can be processed through classical image processing means such as [Chu et al., 2013] describes. Statistical methods extract the average, deviation, skewness and other features of an image are extracted and then combined with historic meteorological data through a standard ANN to predict irradiance. The usage of TSIs is tightly related to clouds, which have a deep impact on how much irradiance reaches the terrestrial surface. Cloud motion is the keystone to the proposal of [Stefferud et al., 2012] where the analysis of multiple TSI images determines cloud motion vectors (direction and velocity of cloud regions) and predicts solar irradiance through these vectors. Other very short-term forecast methods discard motion vectors in favour of advanced pre-processing and feature extraction techniques as is presented in[Fu and Cheng, 2013]. After all sky features have been extracted, a regression model is constructed using image features and ground truth irradiance data. a These approaches are not mutually exclusive as explained by [Bernecker et al., 2014]. In their work both cloud motion and image features are used to predict irradiance using a stream of TSI images.

SVR is a great alternative to ANNs. It offers a key advantage over networks, efficiency. In their work [Zendehboudi et al., 2018] shows that, while both SVR and ANNs show similar results in terms of predictive accuracy, SVRs are highly efficient and less costly in time and resources than ANNs and other machine learning algorithms. The use of SVR in forecasting has been progressively growing over time and many authors have combined it with other techniques. For example, [Li et al., 2016b] combines SVR with ANNs and [Fan et al., 2018b] combines it with XGB.

Other less common Machine Learning algorithms can predict solar irradiance as accurately or better than SVR or ANNs, despite their infrequent use. In their work [Huang et al., 2013] proposes a combination of autoregression with a dynamic system to predict irradiance at hourly scale. Recently deep learning techniques (or deep CNNs) are starting to be applied for forecast-

ing using TSI images. This approach is used by [Siddiqui et al., 2019] who tackles the forecasting problem using video images. They analyse video feeds through a CNN and combines it with more meteorological data. In this case, the video is itself a time series and the images are the individual data points used to forecast irradiance. These complex networks require vast amounts of data resources to develop as [Torres et al., 2019] explains. In their work, multiple data sources (measurements, weather data, plant information...) is combined into a CNN to produce accurate forecasts of next-day photovoltaic production. The review of [Ren et al., 2015] evaluates solar and wind energy forecasting techniques. Hybrid models and ensembles, show a clear tendency to improve predictions when compared against individual machine learning algorithms. Among the evaluated ensemble algorithms, the ANNs used in ensembles tend to offer the best results for the solar forecasting problem.

Time series for Wind Energy Forecasting

As solar irradiance can be used to predict photovoltaic energy, wind energy produced by turbines can be estimated throgh wind velocity. The study from [Lei et al., 2009] reveals the parallels between solar and wind energy, and shows that their similarities are abundant. As in solar energy forecasting, one of the most popular machine learning algorithms are ANNs, which overcome the results estimated by the NWPs for wind energy too.

For wind energy, the SVR algorithm is also used as [Mohandes et al., 2004] indicates. Their study compares ANNs and SVR over the region of Saudi Arabia to predict wind speed. It is shown that the error produced by SVR is smaller than the MLP used. These conclusions are supported by [Zendehboudi et al., 2018] where, with solar energy, the impact of SVR in wind energy is addressed. The use of SVRs in univariate time series is possible as [Santamaría-Bonfil et al., 2016] reports.

Other, less common techniques are the ELMs as it was used by [Wan et al., 2014]. They are applied to forecast the generation of wind energy. It is also possible to employ a genetic algorithm to optimize the internal parameters of the SVR as [Liu et al., 2014] proposes to use in the north of China. Evolutionary algorithms are explored by [Ghorbani et al., 2016] who, on top of an ANN, applies Genetic Expression Programming (GEP) to predict hourly wind speed.

2.3.3 Forecasting by Integration

Contrasting the extended use of time series forecasting, integration (also blending in the literature) has seen less use in prediction tasks over the years. The results of integration tend to surpass any individual method of forecasting, independently of the algorithm used.

Integration has a main disadvantage compared to forecasting: its complexity. Combining multiple forecasts requires already available forecasting models in the first place, which slows down the capabilities of the following steps of the machine learning algorithm. Therefore, the needing of forecasts makes blending slower, more complex and generally harder to implement in an operational setting.

For integration, Equation 2.2 for multivariate time series can be modified to include forecasting models as inputs. Instead of variables V_k, an expression M_k can be included to represent the output of another model. This term has two inputs, (t) that represents the time instant at which the prediction is calculated, and (h) which points to the time instant to be predicted by M_k. Therefore, variations can be made both for (t) as $(t-1)$, $(t-2)$ or $(t-p)$; or for h as (1), (2), (h) or even $(h+1)$. A set of predictions is obtained through the M_k models and the function f can be estimated by using the predictions as inputs to the algorithm.

To compare with the original equation, an example integration model is represented in 2.3, where a function integrates two models M_k that make prediction at different times and for distinct horizons to make a forecast at $P(t+h)$.

$$P(t+h) = f(M_1(t,h), M_2(t,h), M_1(t-1,h), M_2(t,h+1)) \tag{2.3}$$

This representation of M_k can be visualized in Figure 2.2. This figure depicts how a M_k model operates. Three predictions are given in this figure, a prediction on $t-p$ is given for the time instant $t+h-5$, a second prediction is given at $t-5$ to forecast at $t+h$ and a third prediction is given at t for $t+h+5$. These forecasts are built by a common forecasting model.

Related applications to forecast integration in meteorology, wind and solar energy are reviewed in the following subsections.

Integration for Solar Energy Forecasting

The question arises of whether an optimal blending of models can provide enhanced forecasts at any forecasting horizon and station compared with the

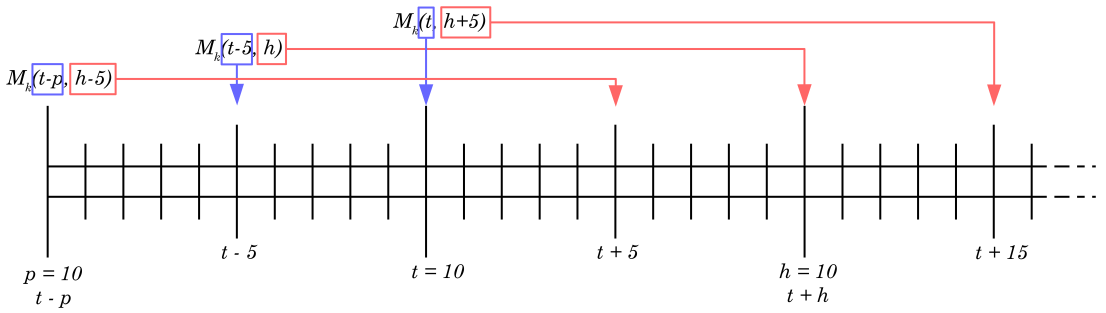

Figure 2.2: Visualization of predictor models M_k.

most accurate forecasts derived from a single model, taking advantage of their synergies and complementarities [Vislocky and Fritsch, 1995]. This question is not new in the framework of weather forecasting research.

Already, early approaches have shown that the blending of different forecasting sources provides enhanced precipitation and temperature forecasts. The production of renewable resources depends directly on meteorological causes and repeated efforts have been made to correctly estimate weather conditions. One early method is proposed by [Xie and Arkin, 1996] who uses the integration approach to predict the amount of monthly rainfall. Machine Learning is not used, instead an statistical method is applied. The algorithm finds the most probable prediction within the average of an area and then combines it with the observations. This algorithm performs adequately, being able to compute the monthly rainfall amount with global coverage. It is also possible to use prediction integration to estimate temperature as [Salazar et al., 2011] proposes for the North America region. Their methodology consists in the combination of multiple Regional Climate Models (RCMs) along with observations from historical data between the years 1971 and 2000. To produce predictions, RCMs are combined with observations through a spatio-Temporal Bayesian model that detects discrepancies between simulations. Results show that the physical models RCM tend to underestimate temperature, a phenomenon that the algorithm corrects.

For renewable energies in general, model blending has been identified as a great possible avenue for improving forecasting accuracy solar irradiance [Tuohy et al., 2015]. Precedents are found in wind energy where the review by [Tascikaraoglu and Uzunoglu, 2014] identifies various challenges of integration. In this review it is concluded that de combination of models is a better approach than using any unique model. However, this has issues, for example that at some prediction horizons the best model is not chosen, or the increase in computational resources, modelling and time. Therefore, the integration of wind energy predictive models shares the typical disadvantages

of any integration technique.

The integration methods proposed across the literature mostly make use of statistical integration of data via linear regression or other estimates as proposed by [Lorenz and Heinemann, 2012]. A weighting of models can drastically improve prediction as it was proposed by [Lorenz et al., 2016], whose results show that linear regression applied to NWP and CMV models can improve the overall accuracy of forecasts on horizons from 0 to 5 hours. The system described in [Haupt et al., 2017] blends several short-term forecasting models by weighting the model contributions. A set of many NWPs used for nowcasting and day-ahead forecasts are blended. For nowcasting it includes the Weather and Research Forecast Model (WRF-Solar), the Cooperative Institute for Research in the Atmosphere Nowcast (CIRACast), the Total Sky Imager Nowcast (TSICast),the Multisensor Advection-Diffusion Nowcast (MADCast) and an statistical nowcasting model (StatCast). For the day ahead forecasts, they use WRF-Solar, NAM, HRRR, the Global Forecast System (GFS), and the Canadian Global Environmental Multiscale Model (GEM). Both forecasting models are blended by linearly combining the predictions by weighting predictions according to their historical performance at each lead time. The overall results of the combined forecast significantly improves the individual predictions. It is concluded that statistical combinations of forecasts are a powerful tool to predict solar irradiance.

Integration of forecast by non-linear physical approaches can also be found. For example CIRACast [Miller et al., 2012, Miller et al., 2018] utilizes winds derived from NWPs to advect cloud field estimates, based on satellite images, to provide short term forecasts. The ability of NWPs to diffuse and advect cloudines information is used by the Multisensor Advection-Diffusion Nowcasting model (MADCast) [Descombes et al., 2014] where cloud retrievals derived from the Multivariate Minimum Residual (MMR) scheme [Auligné, 2014a, Auligné, 2014b] are transported and diffused in three dimensions within the WRF model.

Machine learning is a very popular approach to forecasting and time series in general, and it has seen some use in solar forecasting with ever growing popularity. Previous work shows the potential of using machine learning to combine NWP forecasts. For example, Artificial Neural Networks (ANNs) can be applied as proposed by [Voyant et al., 2012], where a NWP and meteorological current measurements are merged by means of an ANN. In another approach [Lu et al., 2015], three different NWPs are used to forecast meteorological variables, from which predictor variables are obtained and they are blended with RF to predict day-ahead irradiance. The results again show that this

combination is more accurate than the individual NWPs. [Wolff et al., 2016] combines irradiance measurements, satellite and NWP to produce better forecasts of PV production using Support Vector Machines with notable results. The work of [Aguiar et al., 2016] builds an ANN where ground measurements, European Center for Medium-Range Weather Forecasts (ECMWF) and satellite data are blended. The experimental results show a clear increase of skill when multiple information sources are used. Machine learning and integration have been also used in operational environments [Hamann, 2017] with multiple configurations of input NWP models, learning algorithms and data size. Recently, [Dersch et al., 2019] have proposed and evaluated an optimal combination of forecasting models specifically designed for obtaining improved DNI forecasts. Machine learning models can also be integrated via blending as proposed by [Xiao et al., 2015]. A set of physical and machine learning models (Autoregressive Integrated Moving Average and ANNs) are combined. The models proposed across the literature have shown promising results for forecasting integration.

Finally, accurate solar forecast at site (station) level are relevant for plant owners. On the other hand, regional average (or aggregated) forecasts for entire regions are important for transmission system operators (TSOs) [Pierro et al., 2017]. As solar energy increases their share electric systems, an enhanced accuracy of the solar irradiance forecast at grid level are needed in order to manage the electric systems [Renné, 2014].

Chapter 3

Irradiance Estimation

Estimation is the inference of present variables through the analysis of historical data (including past and present information). In the field of solar energy there is interest in estimating solar irradiance because it reduces the costs associated with the acquisition of irradiance data. Therefore, two research challenges are met within this chapter: the automatic cloud classification problem and the image irradiance estimation. The first consists in the automatic labelling of sky types given sky images and ceilometer information. A ceilometer is a device capable of measuring cloud height and thickness. The second research challenge is estimating solar irradiance from sky images using CNNs. This type of network is very effective working with raw images in general. Raw sky images contain a lot of information about irradiance and clouds, so the CNN algorithm can process and extract this information automatically to estimate irradiance.

In this Chapter, two main sections are found, one for each research work. Section 3.1 examines the cloud classification problem, which explains the features and inputs, machine learning method, data, methodology and finally the experimental results and the respective discussion. Section 3.2 studies the irradiance estimation problem where the process of creating new information channels is explained, followed by the CNN architectures designed (single camera and multi-view), the data, methodology, results and discussion.

3.1 Automatic Cloud Classification

To accurately estimate solar irradiance, the cloud classification problem is relevant. Clouds are known to be the main source of interference with solar irradiance [Carslaw et al., 2002]. Knowing the type of currently visible cloud in the sky is highly valuable to estimate irradiance. Different cloud textures, shapes and widths can influence sunlight in known ways [Davies et al., 1975], so ir-

radiance can be estimated accurately when the cloud type is known. Despite its usefulness, the cloud type is rarely used in estimation because an expert is required to label the type. Experts are expensive and cloud classification is a time consuming task, unfit for operational conditions. Additionally there are several instances where multiple types of clouds can simultaneously overlap in the sky, a very hard scenario to classify even for experienced experts [Hoyt, 1978]. There is a necessity for an automatic cloud classification model [Vázquez-Cuervo et al., 2004]. Also skies with multiple cloud types and layers at the same time are rarely considered and accounted for. In a recent work [Wacker et al., 2015] reported this kind of skies to be highly challenging in automatic cloud classification.

A cloud classification model benefits from many sources of information. An unexplored possibility is the ceilometer, which uses lidar detection to measure cloud height and depth. Certain types of clouds appear at higher altitudes and vary greatly in thickness, therefore the data provided by a ceilometer tool can prove valuable. No work in the literature has ever made use of the ceilometer information for cloud classification. Coupled with camera information from sky image statistical features, the accuracy of cloud type classification can be improved with the inclusion of ceilometer inputs. In this thesis, the influence of this new source of information is explored in depth, testing the developed models with hard to detect clouds and ten different cloud classes. Machine learning is used to develop classification models using ceilometer information.

To sum up, the focus of this research is to test whether the ceilometer information has any impact on the classification accuracy of machine learning models.

This work is organized as follows. First the inputs of machine learning are presented in Section 3.1.1 where image features and ceilometer information is described. The random forest algorithm, used for classification, is explained in Section 3.1.2. The data used is described in Section 3.1.3 where the cloud labelling is described. The experimental methodology (data partitioning, metrics and hyper-parameters) is explained in Section 3.1.4. Finally results from this work are shown in Section 3.1.5 and discussed in Section 3.1.6.

3.1.1 Feature Extraction for Automatic Cloud Classification

For cloud classification, TSI images are commonly used. The usual approach for image processing is to extract image features out of the RGB channels. This approach is often limited as features extracted from image colour chan-

nels lack key components of meteorological information sky characteristics. An alternative is to include inputs from meteorological information, such as the ceilometer cloud height and thickness information, which is proposed for this work. The cloud height and thickness is expected to improve cloud classification characteristics because different cloud types appear at different heights and with different thickness.

A mixed set of features has been used here to automatically classify clouds through machine learning. They are divided in two groups, depending of which hardware is used to compute them: image features, extracted from the camera hardware and cloud layer features which are extracted from the ceilometer. A set of 12 commonly used image features for cloud classification have been extracted from the images, for each image there are also 7 measurements of the cloud layers extracted from the ceilometer. The feature list Section at the end of the feature extraction Section contains a list of the features used and how they are calculated, which are explained in detail on the following sections.

Image Features

Most of the image features used are based on [Heinle et al., 2010], and they have been obtained from the red, green, and blue channels of images. These channels are represented using three matrices I^r, I^g, and I^b, for red, green, and blue channels respectively. Each (i, j) location in the matrices corresponds to a pixel in the image, with integer values between 0 and 255. There are several types of image features: spectral features, textural features, and cloud coverage.

The spectral features (Items from 1 to 4 in feature list Section) use the colour matrix I^c exclusively (where c can be r, g, or b), extracting statistical measures directly from it. Some of them are the average and standard deviation from a channel. These are the simplest from the feature set and require very little processing.

The textural features (Items from 5 to 8 in feature list Section) make use of a grey level co-occurrence matrix (GLCM). This is a transformation over one of the colour channels. The result is a g x g matrix, g being the number of grey levels considered in the image. Thus, every element ($p^c_{a,b}$) of the GLCMs (in Equations 3.5 to 3.8) represents the relative frequency of two adjacent pixel values a and b. c represents the colour of the source channel. The value g has been set to its maximum of $g = 256$. GLCMs represent the relative frequency of two pixel values appearing together in the image, at a given offset (in this case, $x_0 = x + 1$, $y_0 = y + 0$, adjacent in the x axis of the image, where (x, y)

and (x_0, y_0) are the two adjacent pixels).

This matrix is commonly used in image analysis for detecting textures in grey images or in given colour channels and are supposed to give information on the spatial distribution of colour, which spectral features are unable to provide. Textures are relevant in the detection of cloud types. In this proposal, four different textures features are used. They measure different properties of the GLCM and are the following: Energy (homogeneity of grey level differences), Entropy (randomness of grey level differences), Contrast (local variation within the gray level matrix), and Homogeneity (similarity of adjacent grey levels within the matrix).

Finally, a cloud coverage statistic (Item 9 in feature list Section) is used in the procedure. To obtain this cloud coverage, first, the original red-green-blue image was converted to hue-saturation-value (HSV) colour space following [Smith, 1978] and [Jayadevan et al., 2015]. Hue describes the colour itself, while saturation denotes the degree of difference between a colour and gray, and value represents the brightness. Saturation ("Sat" in equation 3.9) fits into the range $[0, 1]$, from white, through the grayscale, to the most colourful hue. Cloudy pixels are detected based on a threshold value T = 0.41 for the saturation value. Pixels (i, j) with a saturation greater or equal than this threshold are detected as clear sky; otherwise, the pixels are classified as cloudy. The percentage of sky covered is calculated using the next formula from equation 3.9, where *cp* and *tp* are the amount of cloudy pixels and the total amount of pixels, respectively.

Cloud Layer Inputs

The ceilometer offers height and thickness information about the cloud type (CBH and CPD respectively) that can help discern differences between similar-looking clouds, which would be impossible to recognize otherwise. Layers in cloud formations are numbered in order of distance from the ground. Layer 1 is the closest to the ground, then layers 2 and 3. Given this, seven new machine learning inputs are defined, six for CBH and CPD of each layer plus an extra input indicating how many actual layers (out of three) have been detected (Items from 10 to 12 in feature list Section). The information for each layer is represented as h^l_μ or t^l_μ, to indicate the mean CBH or CPD of layer l ($l = 1, 2$ or 3). There are usually less than 3 layers present in the sky, therefore the variable L is found as the current number of layers found in the sky from 0 to 3. In sum, a total of seven features are derived from the ceilometer to be used in the automatic classification procedure.

Machine learning algorithms require a fixed number of input features. There-

fore, in case the ceilometer returns information of one or two layers only, the missing layers are filled (up to 3) by replicating the information from the closest layer there is a measurement from. In case there are no layers, values are set to an arbitrarily large number, indicating that clouds could not be detected.

Feature List

For the following expressions n is the side size of the image. Therefore n x n are the dimensions of said image.

1. μ^r,μ^b: Blue and red average. (*Image-spectral*)

$$\begin{aligned} \mu^r &= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n} I^r_{i,j} \\ \mu^b &= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n} I^b_{i,j} \end{aligned} \tag{3.1}$$

2. σ^b: Blue deviation. (*Image-spectral*)

$$\sigma^b = \sqrt{\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}(I^b_{i,j} - \mu^b)^2} \tag{3.2}$$

3. γ^b: Blue skewness. (*Image-spectral*)

$$\gamma^b = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}(\frac{I^b_{i,j} - \mu^b}{\sigma^b})^3 \tag{3.3}$$

4. D^{rg}, D^{rb}, D^{bg}: Red green and blue average differentials. (*Image-spectral*)

$$\begin{aligned} D^{rg} &= \mu^r - \mu^g \\ D^{rb} &= \mu^r - \mu^b \\ D^{bg} &= \mu^b - \mu^g \end{aligned} \tag{3.4}$$

5. EN^b: Blue Energy. (*Image-textural*)

$$EN^b = \sum_{j=1}^{n}\sum_{i=1}^{n}[p^b_{a,b}]^2 \tag{3.5}$$

6. ENT^b: Blue Entropy. (*Image-textural*)

$$ENT^b = \sum_{j=1}^{n}\sum_{i=1}^{n} p^b_{a,b} \cdot \log_2 p^b_{a,b} \tag{3.6}$$

7. CON^b: Blue Contrast. (*Image-textural*)

$$CON^b = \sum_{j=1}^{n}\sum_{i=1}^{n} p_{a,b}^b \cdot (a-b)^2 \tag{3.7}$$

8. HOM^b: Blue Homogeneity. (*Image-textural*)

$$HOM^b = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{p_{a,b}^b}{1+|a-b|} \tag{3.8}$$

9. C: Cloud coverage. (*Image-coverage*)

$$\begin{aligned} &Sat_{i,j} > T; T = 0.41 \\ &C = \frac{cp}{tp} \end{aligned} \tag{3.9}$$

10. h_μ^1,h_μ^2,h_μ^3: Average cloud height. (*Ceilometer-height*)

11. t_μ^1,t_μ^2,t_μ^3: Average cloud thickness. (*Ceilometer-thickness*)

12. L: Number of Layers. (*Ceilometer-layers*)

3.1.2 Random Forest

The Random Forest (RF) algorithm presented in [Breiman, 2001] is used here for cloud classification. RF has been reported to be one of the best algorithms for classification [Caruana and Niculescu-Mizil, 2006] and needs no adaptation to work in a multiclass context. This algorithm builds a set of submodels (single classification trees) to form an ensemble that can predict the class of a given instance. Every submodel is an individual decision tree. A simple example of a decision tree can be found in in Figure 3.1.

To classify an instance, the tree is navigated from the root node to a leaf node. Every nonleaf node contains a decision based on an input feature, which will determine the next node to be visited. The tree continues to be navigated through the nodes taking the path that decision nodes determine. Leaf nodes contain labels and, if a leaf node is reached, then the class is determined as the label of the given node.

The RF algorithm constructs multiple different trees from the same training data by means of a double randomization process. First, in order to build each tree, a new data set with the same size as the training data is obtained by sampling with replacement. Second, instead of considering the whole set of

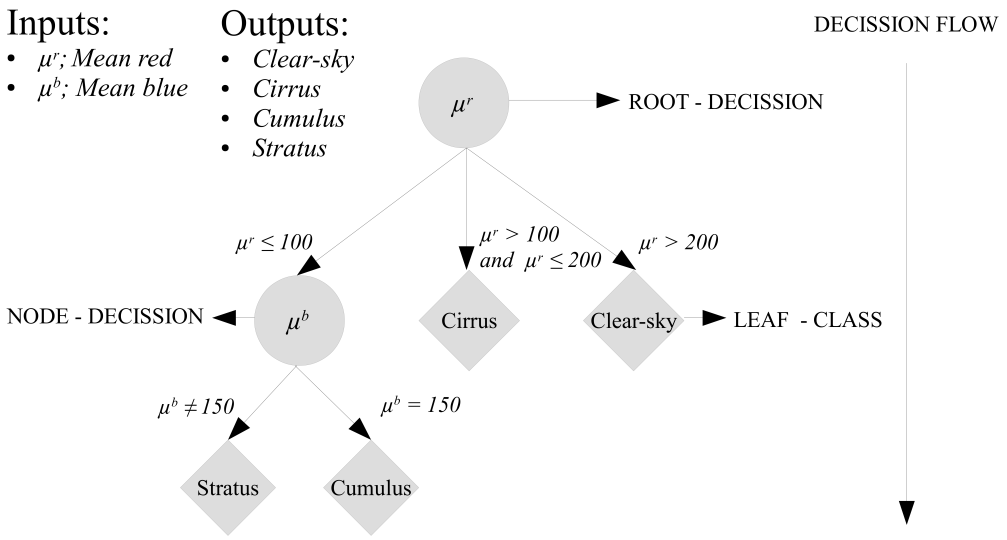

Figure 3.1: Example of decision tree with two input features and four possible classes. It has a maximum depth of 2, 2 decision nodes, and 4 leaf nodes. This tree can classify between clear-Sky, cirrus, cumulus and stratus based on the averages of the blue and red channels.

features, each decision node of each tree uses only a random subset of them (mtry is the parameter name for the size of this random subset, typically much smaller than the whole set of features). The set of decision trees in the RF ensemble classifies new data by majority vote. A diagram of the whole process is represented in Figure 3.2.

Before building the final model, the parameter mtry has to be tuned for optimal classification accuracy. This parameter must be within the range $(1, (F-1))$, where F is the total number of features. The optimal mtry value is obtained by training and testing models with different values and selecting the best performing one. It is important to remark the tuning process uses the training partition only (the test partition is never used for training, parameter tuning being part of that training process). The RF implementation for R has been used [Liaw and Wiener, 2002]. RF has been used together with package caret, which is able to deal automatically with parameter tuning [Kuhn, 2008].

3.1.3 Data set generation

This section explains the dataset used for cloud classification. The camera and ceilometer hardware, data preprocessing procedure and cloud labelling is explained.

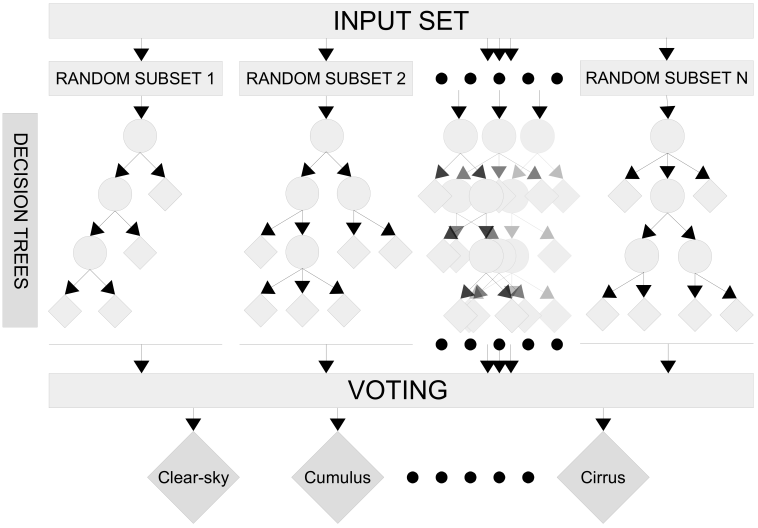

Figure 3.2: Image depicting the construction of a random forest ensemble by random resampling and training of several decision trees. Classification is carried out by majority voting among ensemble members.

Camera and Ceilometer Hardware

All the measurements used were collected at the meteorological station of the University of Jaen, Andalucía (southern Spain), at coordinates 37.7877^{o}N and 3.7782^{o}W, and 454 m above mean sea level.

A total sky imager model Yesdas TSI-880 and a Jenoptik CHM 15k Nimbus ceilometer were installed in September 2012 The TSI-880 is composed by solid-state CCD pointed downward at a hemispheric mirror, which reflects the whole hemisphere (fish-eye vision). Reflection of the Sun is blocked by a dark strip (shadow band), thereby protecting the imager optics. The TSI provides 352 x 288 pixels images every 30 seconds and has been designed for climate/weather applications, showing to be robust regarding environmental conditions [Long and DeLuisi, 1998]. Notably, this camera has been proven to be accurate for the estimation of the cloud cover [Boers et al., 2010, Kreuter et al., 2009, Long et al., 2006, Mannstein et al., 2010]. In the case of high clouds, it is able to report the sky conditions over a spatial domain of about 38 km x 38 km [Mannstein et al., 2010]. In the last years, this sky camera has been used as reference instrument in solar energy applications [Chow et al., 2011, Martínez-Chico et al., 2011, Quesada-Ruiz et al., 2014].

The Jenoptik CHM 15k nimbus ceilometer uses laser pulses at wavelength of 1,064 nm, receiving the backscattered signal over a field of view of 0.45 $fracmrad$. This instrument is able to detect up to five cloud layers simultaneously and to provide their altitude with an accuracy of $\pm$5 m, being its vertical cloud detection range from 5m to 15 km. The sample rate is 15 seconds.

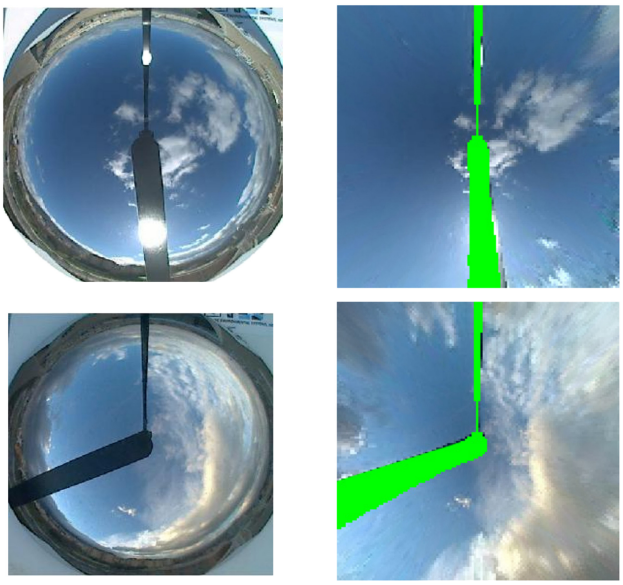

Figure 3.3: Examples of TSI images. Left column represents raw images, right column represents projected images. Top row represent a small cloud sky, bottom row represents a partially overcast sky.

This particular ceilometer is one of the very few ones able to detect clouds above 7.5 km and with less spurious values and better resolution in the upper cloud boundary than other similar instruments ([Boers et al., 2010, Martucci et al., 2010, Wiegner et al., 2014].

Camera and Ceilometer Preprocessing

A total of 717 TSI images, and the corresponding ceilometer estimates, was processed for this study. The images, corresponding to a total of 131 days of the years 2013 to 2015, were selected in order to have a representative sample, with different solar zenithal angles, of the 11 categories described in the following section. Every sample was meant to be representative of 5 minute intervals, that is, images of each of the 11 categories were carefully selected to ensure that during the five previous minutes period exactly the same category was presented. First, the TSI images were masked in order to highlight the border, buildings, and band in the images. Second, the images were projected following [Marquez and Coimbra, 2013]. This procedure transforms the images from a spherical to a rectangular grid. In order to prevent horizon distortion effects, this transformation was conducted only for zenithal angles below 65^o, that is, a 130^o field of view of the camera. Figure 3.3 shows some examples of the TSI raw and processed images.

The ceilometer reports every 15 seconds cloud profiles representative of the column at the ceilometer location, namely, cloud base height (CBH) and

cloud penetration depth (CPD). In this work, up to three different cloud layers were considered. The CPD can be regarded as a proxy of the cloud geometrical thickness. Due to the nature of the clouds (high variability in space and time), ceilometer data should be properly processed in order to provide meaningful information linked to the TSI images, at the 5 min time interval used here. This is particularly relevant for some cloud types, such as cumulus, stratocumulus and cirrocumulus and, in general, cumuliform clouds. These clouds form patches, and therefore, the ceilometer may not report cloud information in some of the 20 samples of the 5 minute evaluation period used here. In addition, the ceilometer sometimes provides spurious measures, or out of range values, which are related to the nature of the backscattering signal processed by these instruments. Given these issues, and since the methodology used in this work aims at emulating a fully operational system, ceilometer data were processed to provide meaningful cloud profile information. First, based on the 20 collected ceilometer samples, a number of candidate group of measurements are selected according to the active layers (up to three). Second, clear-sky values were removed from the 20 samples. Then, based on the CBH values of the remaining samples, a cluster analysis was carried out. The number of centroids in this cluster analysis provided the number of cloud layers, up to a maximum of 3. Finally, for each centroid, a mean CBH and CPD were computed, after applying a filter for outliers. If the 20 measurement are reported as clear sky (i.e., no clouds are detected), the ceilometer procedure final output is the presence of "0" cloud layers.

Since this process has been evaluated trying to mimic an operational system, some problems have been found. Particularly, in about 3% of the samples (24 images), the ceilometer reported no cloud information in cases for which the TSI-880 image was classified in some cloud category different from clear sky. A further analysis confirmed that 15 of these cases corresponded to cirrocumulus and cumulus. These clouds, in many cases, do not cover the whole sky dome and may not overpass the ceilometer column with the 5 minute window here used. The other nine cases correspond to cloud types such as cirrus and nimbostratus. In these cases, the ceilometer was not able to provide the proper cloud information due to technical issues, reporting a very low detection range.

The preprocessing explained above has been carried out by the MATRAS group and is described here in order to facilitate how the images that are used in this work were obtained from the raw TSI images.

Finally, all the features previously described are normalized between 0 an 1 using the maximum and minimum values.

Table 3.1: Dataset summary for each sky class group distribution, average CBH and CPD

Cloud types			CBH	CPD
Seven cloud types + multicloud	10 cloud types + multicloud	Number of images	Mean (SD)	Mean (SD)
Clear sky	Clear sky (CLS)	48		
Cirrus and cirrostratus	Cirrus (ci)	131	9,086 (1,515)	951 (501)
	Cirrostratus (cs)	39	7,684 (676)	1,829 (422)
Cirrocumulus and altocumulus	Cirrocumulus (cc)	13	6,832 (2,023)	469 (238)
	Altocumulus (ac)	75	4,494 (2,257)	726 (516)
Altostratus and stratus	Altostratus (as)	57	6,701 (1,751)	1,858 (607)
	Stratus (st)	53	833 (485)	295 (276)
Stratocumulus	Stratocumulus (sc)	49	1,358 (372)	275 (126)
Cumulus	Cumulus (cu)	54	1,121 (513)	176 (32)
Nimbostratus	Nimbostratus (ns)	42	702 (345)	448 (424)
Multicloud	Multicloud (MC)	156		

Data set

The sky images and the ceilometer information has been classified manually by an expert from the MATRAS group. The 717 samples broken down by class are displayed in Table 3.1. Particularly, two types of labels are identified here. In the first one, up to seven cloud types were used (first column in Table 3.1). These cloud categories are most commonly used in the bibliography [Heinle et al., 2010, Kazantzidis et al., 2012] and try to group cloud types with similar characteristics.

In the second one, compound categories are decomposed into the individual cloud types, resulting in 10 cloud types (second column in Table 3.1). In both cases the multicloud category was added, which aims to represent cases in which the sky is covered by several cloud types at the same time, including the case of several cloud layers. This category is commonly found and should be considered in fully operational systems.

The multicloud category has been scarcely addressed in the literature. [Wacker et al., 2015] describe the problems for automatic recognition of this category, but no attempt for classification was made. Only in [Li et al., 2016a] the multicloud case is considered in an automatic cloud classification procedure. This multicloud category is described as a mix of the sky conditions considered here and covering more than 20% of the sky.

Additional information is included, showing the average cloud height for each cloud type (except clear skies and multiclouds) with the standard deviation in brackets. The cloud thickness is also shown in the same way.

3.1.4 Experimental Methodology

This subsection follows the methods applied to ensure a fair validation. Every experiment shown in Section 3.1.5 has been conducted following the explained methods.

Cross Validation

In order to evaluate the performance of the RF classifier for automatic cloud classification, a cross-validation procedure is carried out. Standard cross validation divides the available data in P equally sized folds or subsets. Then, for every fold p, a model is trained using all folds but p, and tested with fold p (i.e., a performance measure, such as accuracy, is computed for the trained model on fold p). The final cross-validation estimate is the average of the P accuracy values. The standard deviation can also be computed. Common cross-validation practice is followed and P is set to $P = 10$.

However, applying standard cross validation to cloud image data sets can be potentially problematic if the data set contains sequences of temporally related images (cloud images, in this work) taken within short time periods, because some of the images in the sequence might be very similar. This phenomenon is called twinning, and it can lead to optimistically biased cross-validation estimates if very similar images fall into both the training and test partitions. To mitigate this problem, before splitting the data into folds, cloud images are sorted chronologically. Consequently, cloud images that are close in time will most likely fall together either into the training partition or the test partition. This evaluation process avoids the optimistic bias, and it will be more representative of a real situation, because it evaluates the classifier with data belonging to a time period different to that of the training data. However, this stricter validation should be expected to report worse metric values than other state-of-the-art works that use other standard evaluation methodologies.

Metrics

The metrics used for measuring the effectiveness of the models are accuracy and macro-average accuracy. Accuracy is the standard classification success rate:

$$Acc_{Abs} = \frac{S}{N} \tag{3.10}$$

where S is the number of correctly classified instances and N is the total number of instances. The problem with standard accuracy is that classes with more instances have more weight in the success rate. For instance, in an extreme case, if class A contains 95 images and class B contains just 5 images, accuracy is basically informing about class A. In order to measure the behaviour of the model independently of the number of images in each class, macro-average accuracy can be used. This proves a fair comparison for this cloud classification problem due to the imbalance of classes (Table

3.1). Macro-average accuracy is defined as the average of the individual class accuracies.

$$Acc_{Macro} = \frac{1}{M} \sum_{i=1}^{M} \frac{S_i}{N_i} \tag{3.11}$$

where S_i is the number of well-classified instances for class i, N_i is the number of instances for class i, and M is the total number of classes.

Hyper Parameters

All experiments carried out follow the same flow. First, the data set is ordered chronologically and split into 10 different folds. Then, model evaluation is carried out with a tenfold cross validation. In every cross-validation iteration, the training folds are used to select the best mtry parameter value (see Section 3.1.2) and then build the RF model with that value. Then the model is tested with the test fold. Given that RF is a stochastic algorithm, tenfold cross validation has been repeated 10 times, each time with a different random seed (in other words, 10 tenfold cross validations have been carried out). The results obtained are the average of these 10 different runs.

3.1.5 Experimental Results

In this section, results of the different experiments are presented and discussed. One of the aims is to determine the relative contribution of the camera and ceilometer information for cloud classification. Therefore, baseline results were computed by training RF and testing the models using only image features from the camera (spectral, texture, and coverage features). Then, RF models were trained and tested with both camera and ceilometer information. To sum up, eight different experiments were conducted by (1) using up to 7 or up to 10 classes, as described in Section 3.1.3; (2) including/excluding the multicloud category; and (3) including/excluding the ceilometer information. Results of experiments have been organized in two blocks, with 7 and 10 classes (both with and without multicloud), respectively.

7-class Classification Results

Results of the classification procedure when considering seven cloud categories (from Table 3.1), with or without multicloud, are shown in Figure 3.4 and Table 3.2. The accuracy, macro-average accuracy (in percent), and standard deviation (within brackets) are displayed separately for experiments with

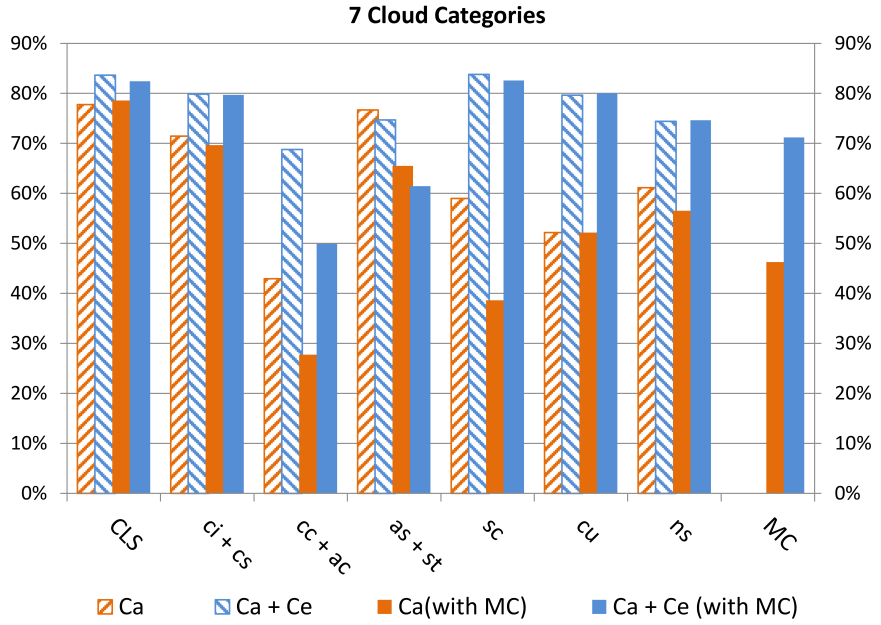

Figure 3.4: Relative frequencies (in percent) of correctly classified cloud classes for the seven cloud types (plus multicloud).

Table 3.2: Overall results for the seven class experiments (plus Multicloud)

		Features used	
	Metric	Ca	Ca + Ce
No Multicloud	Accuracy	64.4% (0.6)	77.3% (0.6)
	Macro-Average	62.3% (0.6)	78.0% (0.6)
Multicloud	Accuracy	55.7% (0.6)	71.7% (0.6)
	Macro-Average	55.1% (0.6)	72.6% (0.6)

camera only (Ca) and with camera and ceilometer (Ca + Ce). In addition, results are presented separately for experiments excluding (seven classes) and including (eight classes) the multicloud category. For the sake of comparison, results excluding and including the multicloud category are displayed separately.

To sum up, results are displayed separately for the four experiments: Using just the camera information (Ca) and both the camera and the ceilometer information (Ca + Ce) but not including the multicloud class, using just the camera information and including the multicloud class (Ca with MC) and using both the camera and the ceilometer information and including the multicloud class (Ca + Ce with MC).

Results clearly show, first, that the use of ceilometer information (Ca + Ce) improves the performance of the classifier for both cases, without multicloud (seven classes) and with multicloud (eight classes). In the former case, the use of the ceilometer improves accuracy and macro-average accuracy by 12.91% (relative 16.7%) and 15.72% (relative 20.16%), respectively (Table 3.2). The improvement is even larger for the multicloud case with 15.97% (relative

22.28%) and 17.46% (relative 24.07%), respectively.

Second, and as expected, including the multicloud class, results in a loss of approximately 5% accuracy when using all the features Ca + Ce, and about 8% when using only the camera (Ca). Interestingly, the ceilometer information allows the classifier to deal better with the extra (and noisy) multicloud class, compared to using only the camera information.

For the experiments without multicloud and breaking down results by cloud type, it can be observed (Figure 3.4) that the ceilometer information increases the accuracy for all cloud types except for stratus and altostratus (in this case, it gets slightly worse by 9.1%). The best improvements are observed for cirrocumulus-altocumulus with 32.7% (relative 47.7%) , cumulus with 30.5% (relative 37.42%), and stratocumulus with 31.6% (relative 37.72%). In the rest of classes, accuracy is also improved to a lesser degree (around 8%). When the multicloud class is included, results are similar regarding the role of the ceilometer. Particularly, the use of ceilometer helps to improve the accuracy of all classes except (again) for the stratus-altostratus class (in this case, accuracy is reduced by just 1.1%). This seems logical, as this class may contain clouds at very different altitudes.

Similarly to the no multicloud case, observed improvements are for cirrocumulus-altocumulus with 21.6% (relative 46.96%), for cumulus with 27.6% (relative 33.85%), for stratocumulus 44% (relative 54.32%), and for multiple cloud type 24.9% (relative 34.03%). For clear-sky, cirrus-cirrostratus, and nimbostratus the improvement is smaller (around 5% and 8%). Finally, the accuracy of the multicloud class prediction is remarkable (around 73%) when using the ceilometer; otherwise, it is just about 48%. The comparison of the results excluding and including the multicloud class reveals some interesting features.

First, when using the ceilometer information, the inclusion of multicloud reduces the accuracy of the classification of just some specific cloud types, namely, cirrocumulus-altocumulus and stratus-altostratus. For the rest of the classes, scores are similar. This result makes sense, since the multicloud category somehow includes the cirrocumulus-altocumulus and the stratus-altostratus classes, which are also composed of several cloud types and cloud layers that can be located at very different altitudes. Therefore, the multicloud type may be confused by these two cloud types. This is what is observed in todo Table 3.3 (the classification contingency matrix). Rows contain the true class, and columns contain RF predictions. Bold entries represent the percentage of well-classified clouds for each cloud type.

Even though the ceilometer helps enormously in the classification, multicloud is misclassified in about 10% of the cases as cirrocumulus-altocumulus

Table 3.3: Contingency matrix results for the seven classes (plus Multicloud) experiment that uses the Ceilometer information

	Label							
Class	CLS	ci + cs	cc + ac	as + st	sc	cu	ns	MC
CLS	**82.2%**	13.9%	0.8%	0%	0%	2.6%	0%	0.4%
ci + cs	2.3%	**79.0%**	2.7%	13.5%	0%	2.0%	0%	0.3%
cc + ac	0%	12.5%	**46.2%**	3.5%	8.9%	1.1%	0.1%	27.8%
as + st	0%	17.6%	1.1%	**64.2%**	0.9%	0.9%	2.9%	12.4%
sc	0%	0%	4%	2.2%	**81%**	2.2%	2.8%	7.8%
cu	9.1%	3.1%	4%	0%	2.2%	**81.7%**	0%	0%
ns	0%	0%	0%	8.4%	6.5%	2.3%	**72.8%**	10%
MC	0%	4.2%	10.6%	10.1%	1.5%	0.1%	0.2%	**73.4%**

Table 3.4: Overall results for the ten class experiments (plus multicloud)

		Features used	
	Metric	Ca	Ca + Ce
No Multicloud	Accuracy	58.8% (0.5)	74.8% (0.7)
	Macro-Average	51.4% (0.5)	66.4% (0.7)
Multicloud	Accuracy	50.6% (0.4)	71.1% (0.6)
	Macro-Average	44.8% (0.6)	63.5% (0.7)

and as stratus-altostratus. Cirrocumulus-altocumulus is classified as multicloud in 28% of the cases. Previous works have also shown that the class cirrocumulus-altocumulus is the most difficult to classify correctly [Kazantzidis et al., 2012, Wacker et al., 2015]. The case of the cirrus-cirrostratus class is different, given that these clouds present a quite similar morphology and, more importantly, are usually located at a very similar elevation. As a consequence, multicloud is misclassified as cirrus-cirrostratus just about 4% of the cases.

To sum up, the performance of the proposed procedure is highly dependent on the ceilometer information. This dependence is particularly relevant for all the "cumuliform" clouds, whose classification accuracy reduces considerably when only the camera information is used. On the other hand, the method showed to be robust against the inclusion of the multicloud class when ceilometer information is used (only the classification accuracy for the cirrocumulus-altocumulus and the stratus-altostratus is reduced).

10-class Classification Results

Table 3.4 shows the results when considering the 10 cloud types displayed in Table 3.1. First, it is observed that the accuracy scores decrease compared to the seven classes results described before. This makes sense, given that the difficulty of classification problems tends to increase with the number of classes. Similarly to the seven class evaluation, a significant increment of the accuracy is obtained when using both camera and ceilometer (Ca + Ce). Overall, these increments are higher than in the seven class case (Table 3.2), indicating that ceilometer information is even more relevant when the number of classes is increased.

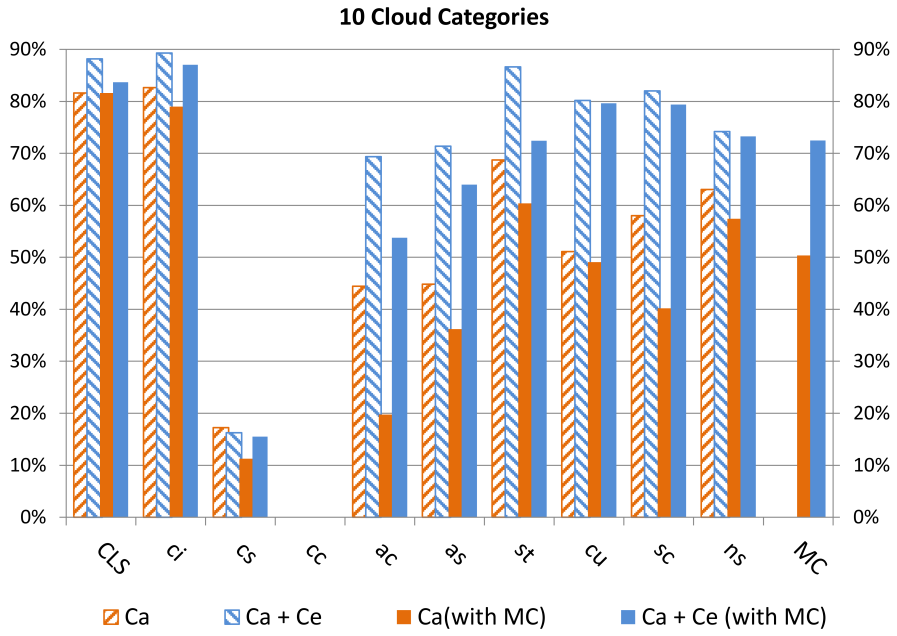

Figure 3.5: Relative frequencies (in percent) of correctly classified cloud classes for the ten cloud types (plus multicloud).

Particularly, the accuracy and macro-average accuracy increase by 20.5% (relative 28.86%) and 18.8% (relative 29.56%), respectively, in the multicloud case. When the multicloud class is included, accuracy and macro-average are reduced by an absolute 3%, approximately, if the ceilometer information is used (if only the camera is used, the reduction is far larger). This result is similar to the seven class experiment. Therefore, the multicloud type does not seem to be an issue in this case. The reduction in the overall performance of the procedure seems to be related with the other categories. Figure 3.5 and Table 3.5 break down results per class. Poor scores can be noticed for the cirrocumulus and cirrostratus classes, which show accuracies of near 0 and 20% respectively, regardless of the use of the ceilometer (Figure 3.5).

Nevertheless, when using the ceilometer information, for some classes (clear-sky) the accuracy increases with respect to the seven class experiment or remain substantially the same (cumulus, stratocumulus, nimbostratus, and multicloud). Regarding the results for the formerly combined classes (cirrocumulus-altocumulus, cirrus-cirrostratus, and stratus-altostratus), now separated, some relevant outcomes were found. For instance, when using the ceilometer and including the multicloud class (Table 3.5), the stratus and altostratus classes show a high accuracy, 78.9% and 64.8%, respectively, higher than the combined stratus-altostratus class in Table 3.3 (62.2%). Note in Figure 3.5 the very relevant information provided by the ceilometer for these two classes. Both kinds of clouds present similar morphological features. The main dif-

Table 3.5: Contingency matrix results for the ten classes (plus multicloud) experiment that uses the ceilometer information

	Label										
Class	CLS	ci	cs	cc	ac	as	st	sc	cu	ns	MC
CLS	**84.3%**	11.8%	0%	0%	0.6%	0%	0%	0%	3.3%	0%	0%
ci	2.9%	**86.4%**	3.1%	0.6%	1.2%	0.8%	0%	0%	2.8%	0%	2.1%
cs	0%	22.5%	**21.5%**	0%	0%	50%	0%	0%	0%	0%	6%
cc	0%	35.7%	0%	**0%**	22.1%	0%	0%	0%	7.1%	0%	35%
ac	0%	11.4%	0%	0.5%	**54.1%**	2.9%	0%	10.4%	0%	0%	20.6%
as	0%	9.5%	14%	0%	2%	**64.8%**	0%	0%	0%	0%	9.8%
st	0%	0%	0%	0%	0%	0%	**78.9%**	1.7%	1.8%	3.9%	13.7%
sc	0%	0%	0%	0%	4%	0%	2%	**80.8%**	2%	3.8%	7.4%
cu	9.6%	2.7%	0%	0%	2.2%	0%	0%	2%	**81.8%**	0%	1.6%
ns	0%	0%	0%	0%	0%	0%	9.5%	7.7%	2.3%	**72.3%**	8.1%
MC	0%	1.9%	2.2%	0.9%	9.0%	5.3%	4.5%	1.7%	0.2%	0.2%	**74.1%**

ference is the location: while stratus are low-level clouds with the CBH below 2 km, the altostratus CBH are typically well above this elevation [Houze Jr, 1993, Kokhanovsky, 2006]. In our case, Table 3.1 data confirms these values, since the mean CBH of the stratus clouds is 833m and the corresponding value for the altostratus is 6,701m. Therefore, it seems that the combined information derived from the camera and especially the ceilometer is able to properly discriminate between these two classes of clouds, even when the multicloud class is included.

The separation of the class cirrus-cirrostratus is not so successful. The cirrus category is reliably classified, reaching 86.4% accuracy (Table 3.5). Note that the information provided by the ceilometer is not highly relevant in this case (Figure 3.5).

Nevertheless, as commented above, cirrostratus results are poor (21.5%). They are classified as altostratus in 50% of the cases and as cirrus in 22.5% of the cases (Table 3.5). These results can be explained based on of the similar characteristics of the altostratus and cirrostratus clouds. Particularly, the altostratus clouds present a mean CBH of 6701m, with 1,701m standard deviation value (Table 3.1). The corresponding values of the cirrostratus clouds are 7684 m and 676 m. These experimental values are confirmed in the bibliography, which states that the range of elevation in middle latitudes is 2–7km for the altostratus and 5–13km for cirrostratus [Houze Jr, 1993, Kokhanovsky, 2006]. The CPD for both types of clouds is also similar: 1,858m and 1,829m for the altostratus and cirrostratus, respectively (Table 3.1). Therefore, these clouds cannot be discriminated just based on the CBH and the CPD. The main difference between these two kinds of clouds is the usual presence of the halo feature in the cirrostratus clouds but not in the altostratus. This particular feature seems to be not resolved by the image characteristics used here.

Finally, the poorest results are obtained for the cirrocumulus-altocumulus discrimination. Particularly, cirrocumulus clouds are systematically misclassi-

fied as altocumulus, cirrus, cumulus, or multicloud (Table 3.5). These poor results can be explained based on several reasons: first, because the mean and the standard deviation CBHs values (Table 3.1) of the cirrocumulus (6,833m and 2,023m), altocumulus (4,494m and 2,257m), and cirrus (9,086m and 1,515m) do not allow the use of the CBH to discriminate the cirrocumulus from the other two cloud types. Reference values of the CBHs in middle latitudes are 5–13km for cirrocumulus and 7–10km for cirrus 2–6 for altocumulus [Houze Jr, 1993, Kokhanovsky, 2006], therefore confirming our results. Similar inferences can be derived for the role of the CPDs, which show mean and standard deviation values (Table 3.1) that makes the CPD inadequate to discriminate between these tree cloud types. Again, reference values in the bibliography confirm these findings; particularly, [Houze Jr, 1993, Kokhanovsky, 2006] report the geometrical thickness of the cirrocumulus to be in the range 0.2–0.4 km, which overlaps the thickness of the altocumulus (0.2–0.7 km) and cirrus (0.1–3 km). Therefore, ceilometer information seems not to be relevant to distinguish the cirrocumulus clouds from many other classes. Regarding the sky camera information, from the morphological point of view, cirrocumulus and altocumulus are similar. In addition, cirrocumulus clouds often occur in small sheets located very high in the atmosphere (even 9 km values can be found in the experimental data set here used). As a consequence, and probably also because of the low resolution of the TSI images, the camera is not able to provide distinctive statistics values for this particular cloud class. Altocumulus results are more encouraging (accuracy 54%), although they are misclassified as multicloud in 20.6% of the cases.

3.1.6 Summary of experimental conclusions

In this section a cloud classification model has been proposed with several interesting experimental results. The most important insight from this work is that for every possible case studied, ceilometer information positively impacts the automatic classification accuracy of cloud types. Two different cloud classifications are tested, a common 7-class labelling and a more strict 10-class labelling. These show that, independently of the complexity of the labels, ceilometer information improves classification.

Results have shown that the inclusion of the multicloud type negatively impacts the classification of most cloud types. In this scenario, ceilometer information mitigates the negative impact. This behaviour is repeated when the 10-cloud types are tested, where there is a negative impact on accuracy across all classes, but ceilometer information improves camera results nonetheless.

Two metrics are used for evaluation, accuracy and macro-average accuracy. Results on macro-average accuracy show to be lower than accuracy on the 10-cloud types, this happens because some cloud types are over-represented which could be a source of accuracy loss. This does not happen on the 7-cloud types are evaluated as accuracy across classes remains similar.

Some cloud types have very specific behaviours in relation to cloud classification. Some cloud types improve dramatically with ceilometer information, completely negating the impact of multicloud, such as cumulus, stratocumulus and nimbostratus. Other classes are very hard to classify, such as cirrostratus and cirrocumulus, for which the ceilometer has no impact whatsoever. Given the diverse nature of multicloud skies, when included most cloud types are wrongly detected as such.

Overall, it is concluded that ceilometer information is a valuable source of information that can greatly enhance automatic cloud classification.

3.2 Irradiance Estimation with CNN

Solar energy measuring devices, such as pyrheliometers and pyranometers, are expensive and require constant maintenance, supervision and a high upfront cost, properties that makes acquiring irradiance data inaccessible for most research purposes [Coppolino, 1994]. Sky cameras do not have this problem, as they do not need a large investment or expert maintenance. Estimating irradiance can be difficult due to cloud interferences, however by exploring more advanced machine learning techniques the camera could prove enough to estimate solar irradiance. Sky camera images contain valuable information about the sky type, solar hour, sky brightness, etc, that an algorithm could automatically detect.

In the field of image recognition Convolutional Neural Networks (CNN) have trended in use due to their powerful classification capabilities and automatic feature detection, so they can be considered an adequate technique for irradiance estimations from raw sky images. CNNs have been used for solar energy tasks, but no application to irradiance estimation has been found across the literature.

Cited works in Section 2.2.2 rarely approach the problem of irradiance estimation through automatic image processing, there is always a heavy human expert component when analysing images. When these images are processed automatically, other works are limited to a single fixed camera point. Some make use of temporal combination of multiple images of video feed, but an analysis of various sky images for irradiance estimation has never been tested.

As mentioned before, cloud height and thickness is extremely relevant to solar irradiance and information about them is useful. However a single bi-dimensional image is unable to contain three dimensional information. The usage of multiple perspectives to infer 3D shapes has never been applied to the solar estimation problem, and CNNs can be modified to meet this approach. There have been attempts to merge information of multiple images through convolutional networks [Su et al., 2015] but no attempt has been made for irradiance estimation or forecasting.

Summarizing all the above, the aim of this part of the thesis is to deal with irradiance estimation TSI from images, but using CNNs in order to go beyond merely using statistical features (see Section 2.2.2). Therefore, in this thesis it is intended to study whether CNNs can obtain more accurate irradiance estimations. In order to do that, two adaptations to the standard CNN architecture and use have been proposed. First, instead of a single image, three cameras that record images from different points of view will be used. A modification of the standard CNN architecture has to be proposed, so that three images can be successfully combined. In this regard, two CNN architectures are proposed, a single camera model, and a multi-view model capable of processing and combining multiple sky images into a single estimate. Second, instead of using only the RGB channels which are commonly used in image processing, two additional channels relevant to irradiance estimation are added: a channel that contains information about which pixels are cloudy and another channel that tells the distance of each pixel to the sun centre.

The structure of the following section is presented. The CNN architectures are described in detail in Section 3.2.1, where a typical CNN, the single camera CNN and two multi-view approaches are proposed. The data set used is described afterwards in Section 3.2.2. The methodology used in the experimentation is detailed in Section 3.2.3 including the baseline methods used. Finally the results and discussion of this experimentation are presented in Sections 3.2.4 and 3.2.5 respectively.

3.2.1 Single and Multi-view models using CNNs

CNNs are commonly applied to classification in image recognition. However, by altering the output neurons to return numerical values it can easily be adapted to a regression problem such as irradiance estimation. All-sky images contain vast information about the sky condition and by analysing the colour channels, irradiance may be inferred. Before the proposals of this thesis are presented, a standard CNN architecture with the common operators is ex-

plained, followed by the single camera architecture and two alternatives of the multi-view architecture.

Standard CNN architecture and operators

A CNN works by applying several operations over the input channels. The basic operation is the convolution. If an image has nxn dimensions, a convolutional filter of side k (kxk dimensions) where $k \leq n$ can be applied. The filter is a matrix of values $K_{i,j}$, that acts as weights to be learned throughout the learning process. The filter is moved over the channel with a given step size s. The result of applying a convolutional filter is a new matrix of side $\frac{n-k+1}{s}$ pixels. Therefore, a convolved pixel $I'_{a,b}$ of an image I with m channels from a filter K is the result of following Equation 3.12.

$$\begin{aligned} I'_{a,b} &= \theta + \sum_{c=1}^{m} \sum_{i=a'}^{k+a'} \sum_{j=b'}^{k+b'} K^c_{i-a',j-b'} \cdot I^c_{i,j} \\ a' &= a \cdot s \\ b' &= b \cdot s \end{aligned} \tag{3.12}$$

where a and b are the positions in the result I' matrix. Here two matrices are used K and I, which represent the convolutional filter and image, respectively. The term $I^c_{i,j}$ represents the pixel at position i, j from the image channel c. The step s offsets pixel positions, therefore a' and b' represent the adjusted positions to the step size. The maximum values of a and b are contained within the set $\{x \in N | x < \frac{n-k+1}{s}\}$. The constant θ is the bias value (the concept is the same as in standard multilayer perceptrons).

The convolutional operation can be a dilated convolution, where the filter is filled with zeroes in-between values to increase the filters size without increasing the number of weights in the filter (see Figure 3.6). A filter position $K_{i,j}$ dilated by d would end up in the dilated filter position $K'_{a,b}$ with the same value where $a = i \cdot d$ and $b = j \cdot d$. Every other position in K' is filled with zeroes. The filter K' can be used in place of the regular K filter to cover bigger regions of a matrix without compromising memory and speed of the network. The result of filling a filter K with sides kxk using a dilation of d will have dimensions equal to $k' = k + 2d$. The effects of transforming K to K' depending on the d value are represented in Figure 3.6. This figure shows the effects of dilation on a 3x3 kernel, where it is dilated by $k = 0$ (or not dilated), by $k = 1$ which separates the original values by a row/column of 0 values to get a kernel with dimension 5x5 padded with zeroes, and by $k = 3$, where the original values are all separated by three 0 values.

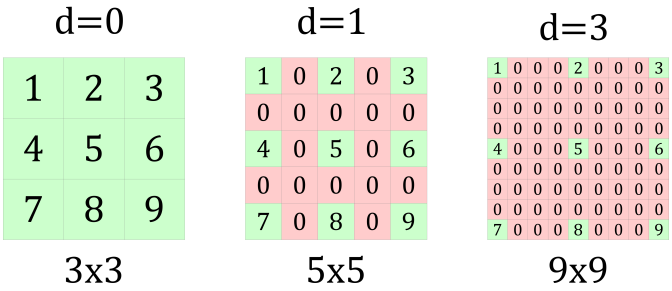

Figure 3.6: Effects of dilation on a typical convolutional kernel. Left figure: 3x3 kernel dilated by $k = 0$ positions. Center image: 3x3 kernel dilated by $k = 1$ position. Right image: 3x3 kernel dilated by $k = 3$ positions.

Usually, CNNs use more than a single filter, resulting in a set of matrices (or convolved images) instead of a single one. These convolved images are named as $I^{(1)}_{a,b}$, $I^{(2)}_{a,b}$, $I^{(3)}_{a,b}$... up to the number of filters $K^{(1)}_{i,j}$, $K^{(2)}_{i,j}$, $K^{(3)}_{i,j}$... used.

However, convolutions are a linear operation and to achieve non-linearity, a transformation of the outputs from linear to non-linear is required. A common approach is applying the Rectified Linear Unit function (RElu) shown in Equation 3.13. This rectification is applied to every matrix position $I'_{a,b}$.

$$f(x) = \max(0, x) \tag{3.13}$$

The last major operator in a convolutional network is the pooling. In a similar way as the convolution, a window of size k and a step size s (most times higher than 1) is chosen and applied over the matrix. The pooling operator applies a function to the window without the use of a filter. Common pooling operations compute the maximum or average of the given window. Here, the average pooling operation has been chosen, as described in Equation 3.14.

$$\begin{aligned} I'^{c}_{a,b} &= \frac{1}{k^2} \sum_{i=a'}^{k+a'} \sum_{j=b'}^{k+b'} I^{c}_{i,j} \\ a' &= a \cdot s \\ b' &= b \cdot s \end{aligned} \tag{3.14}$$

Typically, once several convolutional and pooling layers have been applied, a Multi Layer Perceptron (MLP) is used to perform a final processing. The inputs the MLP layers (known as fully connected) are the outputs from the sequence of convolution, rectification and pooling operations. These outputs are a set of bi-dimensional channels, which are flattened into a single unidimensional vector. Fully connected layers have as many weights as connections between the input vector and neurons. Then, the neuron e with

vector of weights W^e transforms the vector of inputs. The output of channel flattening and weight product for any neuron e is shown Equation 3.15.

$$Y_a^e = \theta^e + \sum_{c=1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{n} I_{i,j}^c \cdot W_{i+nj}^e \tag{3.15}$$

where I_{ij}^c are the bidimensional matrices resulting from the convolutional/pooling sequence. The outputs obtained from the fully connected layer can also be transformed to non-linear values through the ReLU function.

Due to the high number of neurons and weights, the whole convolutional network is prone to overfitting. An operator that prevents this is the dropout. This technique disables a random proportion of neurons (or filters) from a layer. During a cycle, the set of disabled neurons are disregarded, and their activations are not carried over to future layers.

The weights and biases of the filters and connected neurons have to be optimized. There are many optimizers such as Stochastic Gradient Descent (SGD), Adagrad, Adam, etc from which the Adam optimizer [Kingma and Ba, 2014] has been found to be the most suitable. It has been chosen because it has shown a good behaviour in many tasks and does not require as much tuning as SGD.

An example of a typical CNN architecture is shown in Figure 3.7. A RGB image is the input of this network, processed through a set of layers. Each layer contains a convolution, an activation and pooling operations. The result from the layers are flattened and processed through a fully connected layer, and an output is given.

Single image model

CNNs could be directly applied to estimate solar irradiance. However, due to the difficult nature of this problem, some key modifications to the standard CNN presented in the previous section have been performed. The most inmediate approach to a single camera model, would be using the standard three RGB channels. However, by nature, the irradiance estimation problem requires additional information about the sky as preliminary results had shown. Therefore, two additional channels are proposed, containing domain-specific information about the sky condition. This added channels contain contextual information that, although present in the camera, is not explicit within the RGB colour channels.

The additional channels and an example of an original image are shown in Figure 3.8. First the the cloudiness of the image is considered. This prop-

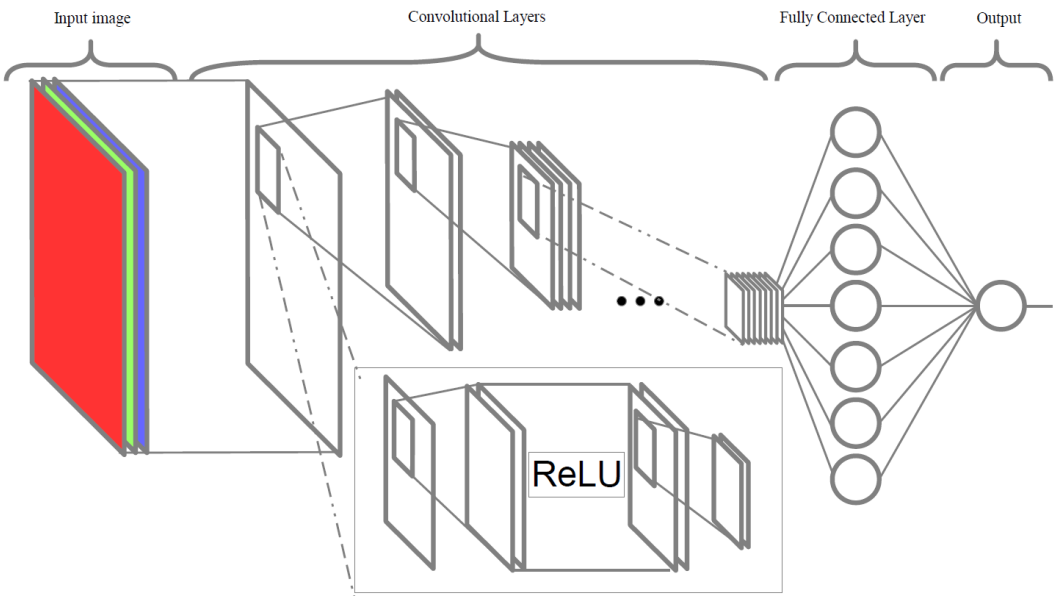

Figure 3.7: An example of a standard convolutional neural network architecture.

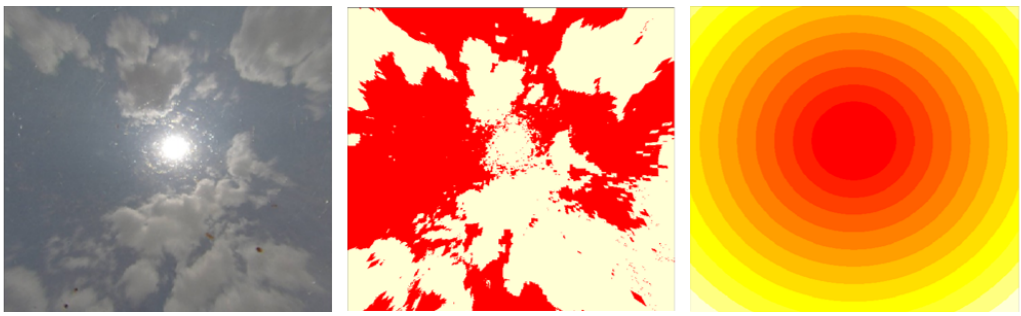

Figure 3.8: Examples of an original image and additional channels. Left image: Original image of a partially overcast sky, composed of RGB channels. Center image: Hyta filter, bright pixels are clouds (1), dark red pixels are clear sky (0). Right image: Sun distance filter, brighter pixels are further away from the sun disc centre.

erty of sky pixels can be directly inferred from RGB channels via filtering. A more advanced pixel cloud detection than in Section 3.1.1 has been used in this work, with the HYTA [Li et al., 2011] algorithm. HYTA thresholding uses two methods to detect cloud pixels, a fixed thresholding algorithm, and an adaptable threshold based on histogram readings. Depending on the B/R channel ratio of the image, the first or second method is used. The end result from this process is a 500x500 channel of 1s and 0s that represent whether a cloud has been detected or not respectively at each of the pixels. This enhanced cloud detection has been used on the RGB channels of the transformed images to extract a fourth channel containing a description of the cloudy pixels present in the image. This is due to HYTA being more accurate at cloud detection than a simple thresholding in general and particularly for images without sun band.

Using information from the sun position and radius, another channel has been added. Knowing the sun centre pixel of the projected image and the radius of the sun disc, the distance from the sun for each pixel has been calculated. Considering a tuple of coordinates (x_c,y_c) representing coordinates x and y of the centre of the sun, and a second tuple of coordinates (x_p,y_p) representing the desired pixel coordinate to be measured, the euclidean distance is calculated using Equation 3.16. This information is computed for each pixel of the image to get a fifth channel of 500x500 where pixels inside the radius of the sun centre are 0 and pixels outside the sun disc increase in value as the distance increases according to the euclidean distance.

The reason for including this new channel is that the meaning of the brightness of a pixel may depend on whether the pixel belongs to a bright region in the image (or not) [Alonso-Montesinos and Batlles, 2015], and the distance from the sun may be a proxy for this information.

$$z = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2} \tag{3.16}$$

In addition to the new information channels, another modification is introduced to the standard CNN. A dilated convolution is set as the first convolutional layer. This operator is different from the other convolutional layers described as it covers a bigger area with the dilated window and excludes the average pooling operator. This modification is introduced due to the complex nature of TSI images as a wider kernel coverage may detect hidden textures within any channel.

Although not uncommon, average pooling is far less used than the maximum pooling operator. In this architecture, the average pooling operation is used because the irradiance of any image region could be expressed as the sum of the local irradiances from the individual pixels.

This application of CNNs with the inclusion of the previously mentioned additions and modifications is represented in Figure 3.9

Averaging the single camera model for multi-view estimation

For this work, images from three different cameras are available, with an associated irradiance measurement. Further information can be inferred from this set of images and their relationships.

All multi-view models will use multiple images with their respective channels to estimate irradiance. An immediate approach is to reuse the single camera model as the basis for a direct multi-view approach. The Average CNN follows two steps. First, all three simultaneous images follow the single cam-

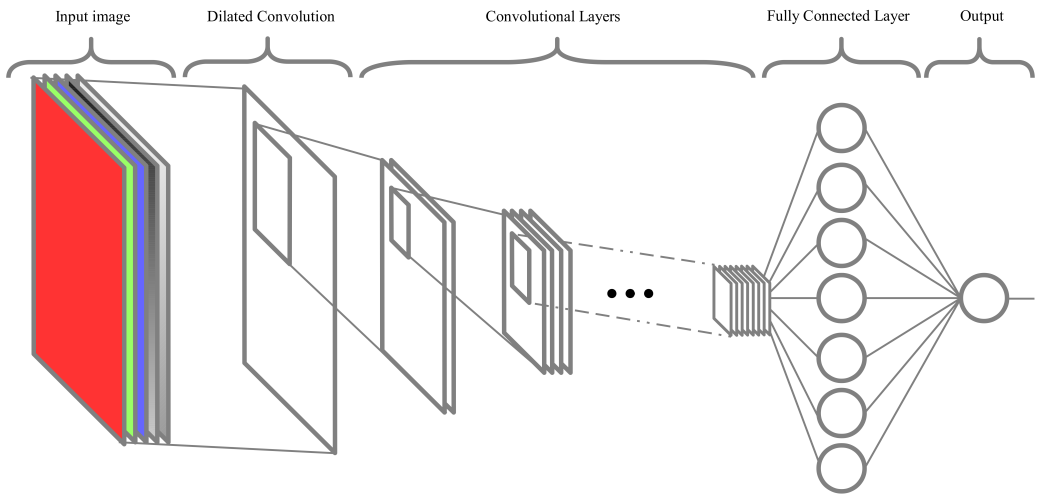

Figure 3.9: Representation of the single image CNN model

era model architecture up to the final convolutional layer. After these, the outputs from the three CNNs (a CNN trained for each image) are averaged and flattened for the fully connected layer to process.

As stated, this is an immediate adaptation of the single camera model to combine the information from three cameras, and the utility of this method is later tested.

Multi-view model

However, the averaging of the CNN outputs might be too simple to detect further synergies in the image set, so it is hypothesized that the CNN may be able to find complex relationships within the image if the combination of outputs happens in the middle of the CNN flow. Therefore the multi-view CNN model is proposed.

The CNN architecture proposed in this work for multi-view can be seen in Figure 3.10. Every image is connected to a dilated convolutional and activation operator, which in turn is connected to a convolutional network containing three convolutional layers. These initial operations are identical for the three images, sharing the exact same filter weights and biases between image pipelines. The end result from the three convolutional networks are a set of filtered images, which have gone through the same identical filters and operators. The outputs from this first CNN_1 are noted as $I_{a,b}^{(c)}$, for example $I_{a,b}^{1(c)}$ would be the output from the first image, and can be 1, 2 or 3 for each output (one per image).

Every filtered matrix has a counterpart in the other outputs, so their in-

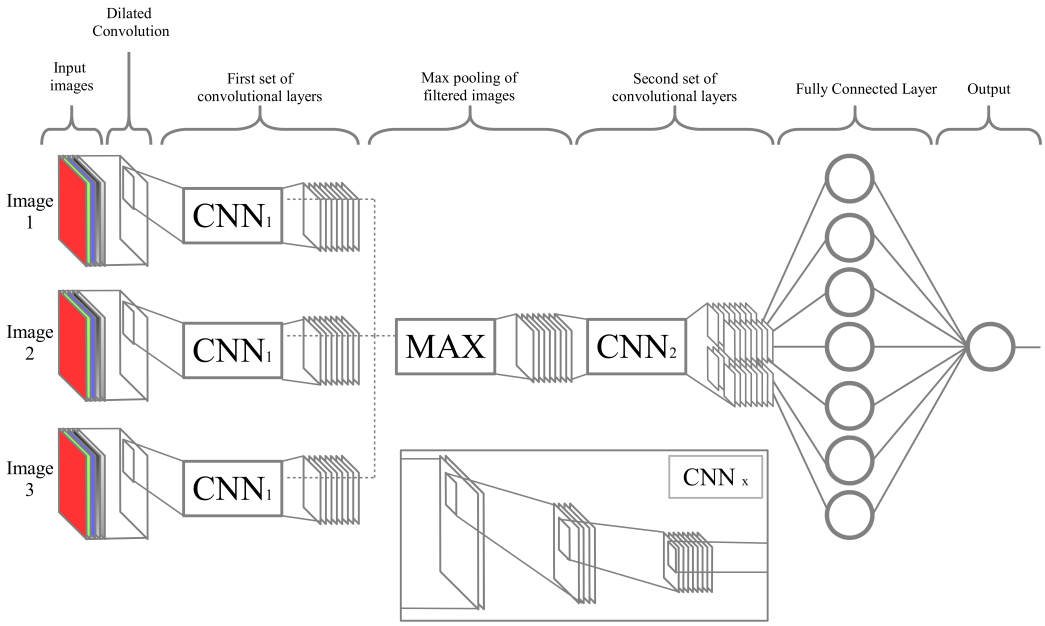

Figure 3.10: Representation of the multi-view CNN model

formation can be merged through a pooling operation presented in Equation 3.17. This pooling operation combines each pixel as the maximum from the three $I_{a,b}^{(c)}$ convolution outputs, resulting in a new $I_{a,b}^{\prime(c)}$ combination of the outputs.

$$I_{a,b}^{\prime(c)} = \max(I_{a,b}^{1(c)}, I_{a,b}^{2(c)}, I_{a,b}^{3(c)}) \tag{3.17}$$

where $I_{a,b}^{\prime(c)}$ is the resulting c-th matrix value at a, b and $I_{a,b}^{1(c)}$ is the input value from the first set of values, obtained from the first image. The max pooling creates the new I' that can be put through another set of convolutional operators in the same way as before. After the information is combined, it is processed through a second set of convolutional layers CNN_2. The rest of the network is an standard fully connected layer with ReLU activation and dropout.

3.2.2 Data set generation

Though similar in shape, the dataset used for irradiance estimation is different from the one previously described in Section 3.1.3 for automatic cloud classification. In this case, more images and irradiance information have been gathered from a diverse set of spatially separated points. The following sections describe the characteristics of the data set used in this section, and the preprocessing that has been carried out.

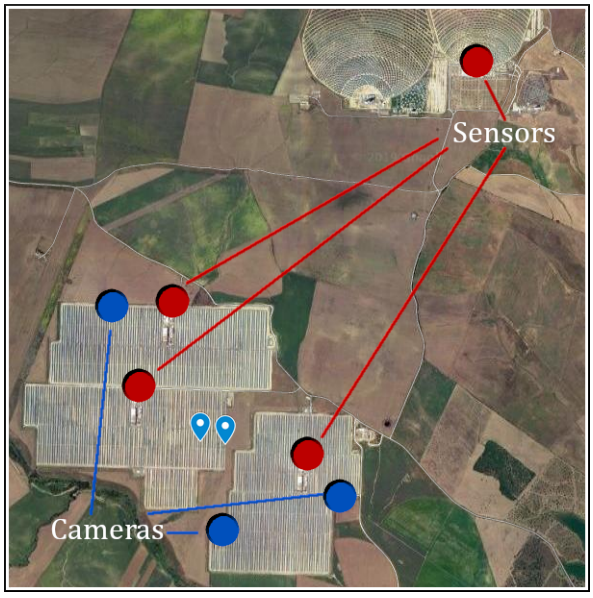

Figure 3.11: Spatial representation of the cameras and sensors. Blue dots represent the cameras and red dots represent the irradiance sensors.

Image and irradiance measure gathering.

Data has been gathered in a similar way as in Section 3.1.3, using TSI sky cameras and irradiance sensors. Measurements used have been gathered from three data stations at the geographical region of Seville, Spain.

A commercial-grade AXIS M30007-PV Network Camera with a fish-eye lens has been used. Each camera produces images with 1944x2592 pixels, however the useful information from the sky is contained within a centered circle of approximately 972 pixels. Two devices are placed in each given point for irradiance measure gathering: A Kipp&Zonen CHP-1 pyrheliometer, for DNI information and a Kipp&Zonen CMP-6 pyranometer, for GHI measurements.

All cameras and sensors are distributed spatially as represented in Figure 3.11. There are three cameras at different positions, the first camera is positioned at coordinates 37.409922^{o}N and 6.272902^{o}W,, the second at 37.412107^{o}N, 6.262683^{o}W and the third at 37.4258^{o}N, 6.282963^{o}W all three at 35m above sea level.

To measure the irradiance of this region accurately four irradiance sensors are placed at coordinates: the first one at 37.414261^{o}N, 6.265571^{o}W, the second one at 37.419067^{o}N, 6.280442^{o}W, the third at 37.42482^{o}N, 6.27761^{o}W all three at 55m above sea level; and a fourth sensor at 37.442011^{o}N, 6.250549^{o}W at 44m above sea level.

Preprocessing

The processing of images is discussed in Section 3.2.1, where the additional channels are described.

Following the quality control measures proposed by [Long and Dutton, 2010], a filtering has been passed through irradiance data. Some filters require DHI, not available, so the filters that make use of DHI have been ignored. On general terms, the quality control tests are passed to avoid errors in measurements and exclude extreme outliers.

Four measurements of irradiance are taken simultaneously for each time instant. To obtain the real irradiance value, the four measures are averaged. The slight errors made by instruments are cancelled out when the average is taken, producing more accurate results.

The images have been preprocessed to fit a square image following the same procedure as in Section 3.1.3 to a resolution of 500x500 pixel square. The flattened image contains three colour channels, red, green, and blue. Images are normalized from 0 to 1 using the maximum and minimum values from each channel of the complete set of images.

A common technique in solar irradiance estimation is predicting the clear sky index K_t value from a irradiance value. The K_t is a ratio between two values, the real irradiance and the clear-sky (or extraterrestrial) irradiance. The latter can be interpreted as the maximum possible irradiance at a particular time and place. This ratio proves useful because the influence of solar time is removed from irradiance as this information is not implicit. This simplification can be made because the clear-sky irradiance can be calculated analytically out of already given information (latitude, longitude, time, etc.).

The irradiance information is also normalized from 0 to 1 using minimum and maximum values from the averaged irradiance measurements.

Data set

Described as before, the dataset counts with two distinct parts, the images and irradiance data. All images come in sets of three, an image from each camera, all taken at the same time. Also the irradiance information is taken at the exact same time, averaged from the last 15 minutes. The irradiance sensors and cameras have gathered data from June to October of 2015.

An example of an image and the additional channels are represented in Figure 3.8. Each image presents five different matrices of 500x500 pixels labelled as red, green, blue, clouds and sun distance.

In addition to the images and irradiance information each instance in the

Table 3.6: Dataset summary for each sky class group distribution, average GHI with absolute (GHI-Value) and clear sky (GHI-K_t) values. Standard deviation is shown for each group in brackets

Cloud Type	Number of images	GHI-Value	GHI-K_t
Clear sky	14	737.91 (119.07)	0.938 (0.019)
Cirrus	100	783.82 (91.93)	0.933 (0.094)
Cirrostratus	100	545.54 (113.12)	0.689 (0.134)
Cirrocumulus	30	578.22 (151.80)	0.686 (0.196)
Altocumulus	122	600.15 (184.42)	0.686 (0.2)
Altostratus	41	248.94 (44.69)	0.295 (0.043)
Stratus	163	180.86 (95.44)	0.224 (0.115)
Stratocumulus	220	615.92 (213.45)	0.738 (0.23)
Cumulus	100	757.53 (112.9)	0.834 (0.128)
Nimbostratus	6	51.71 (26.45)	0.074 (0.037)
Multicloud	263	516.43 (235.94)	0.577 (0.258)
Aerosol	33	592.46 (105.1)	0.633 (0.106)

dataset has an associated cloud type. Added to the previous classification, aerosol images have been included, which represents another difficult type of sky, a clear sky with low irradiance. It is very important to note that the cloud type is only used for data partitioning in the estimation problem. Cloud type is very representative of the global irradiance, therefore splitting the dataset by cloud type will guarantee a fair representation of different sky types for the estimation of irradiance. Additionally the cloud type allows for further analysis in the experimentation section by analysing the regression error metrics class by class and performing their macro-average. Table 3.6 describes the dataset by cloud type, showing information about the number of instances and average irradiance values for each sky classification.

3.2.3 Experimental Methodology

To ensure the fair evaluation of the proposed models, the following experimental methodology is used for each experiment. Also the baseline methods and metrics are explained.

Validation

To validate the CNN, the dataset is split in three different subsets: a training set, used for training the CNN; a test set, to study the generalization of the algorithm; and a validation set, to detect and prevent the overfitting of the network. The CNN trains over a set number of epochs, building different models by adapting the values of filters and weights. The decision of which epoch has the best model is an integral part of the training process, which is decided by measuring the validation set rMAE metrics. The epoch at which the model has the lowest validation error is chosen as the final model for that set. The three partitions are made from the initial dataset of 1290 instances

which is divided into sets containing 70%, 15% and 15% of the data for training, validation and testing, respectively.

To ensure a fair representation of sky images and maintain the diversity of the original dataset, the partitions are made obeying to the proportion of cloud classes. This means that train, validation, and test contain the same proportion of instances, according to cloud type. Some under-represented cloud types such as nimbostratus require this partition criteria for the CNN to detect the specific features of that type of sky.

All results shown are validated on the test set and shown in the respective sections.

Metrics

To measure the quality of the blending models, the metrics Normalized Root Mean Square Error (rRMSE) and Normalized Mean Absolute Error (rMAE) have been used. Both are linear transformations of RMSE and MAE relative to the average measured value. RMSE penalizes outliers due to the square transformation of the error and is deemed standard error representation in the meteorology field. However, the MAE, while not being as strict with large errors, is frequently used in operational environments to calculate losses, therefore it also is analysed. The relative transformation allows the error measurements to be independent of the average irradiance and can be interpreted as a percentage. Equations 3.18 and 3.19 represent the rRMSE and rMAE respectively.

$$RMSE(P, D) = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(P_i - D_i)^2} \tag{3.18}$$

$$rRMSE(P, D) = \frac{RMSE(P, D)}{\frac{1}{N}\sum_{i=1}^{n} D_i}$$

$$MAE(P, D) = \frac{1}{N}\sum_{i=1}^{n}|P_i - D_i| \tag{3.19}$$

$$rMAE(P, D) = \frac{MAE(P, D)}{\frac{1}{N}\sum_{i=1}^{n} D_i}$$

where P_i is the i-th prediction for the dataset which is compared with the D_i desired output. The study of irradiance by cloud type can provide more

detailed information about the problem at hand, so the metrics are calculated by class and then averaged as a macro-average. This is also useful because some sky types are underrepresented and the global metrics may be biased towards the majority sky types. Both macro-average metrics can be found in Equation 3.20

$$rRMSE(P,D)_{Macro} = \frac{1}{M}\sum_{j=1}^{M} rRMSE(P^j, D^j)$$
$$rMAE(P,D)_{Macro} = \frac{1}{M}\sum_{j=1}^{M} rMAE(P^j, D^j) \tag{3.20}$$

where P^j and D^j are vectors of predictions and observations from sky type j where $P^j \subset P$ (and $D^j \subset D$). The number of sky types is represented by M.

Hyper Parameters

CNNs have a very large number of hyper-parameters to adjust namely convolutional layers, filters, epochs, etc. First, the number of epochs have been decided by leaving the network a large number of cycles to train. Every time a new minimum validation rMAE value is achieved, the responsible model is saved. Then, when all the epochs are complete the model with the lowest rMAE on validation is chosen and the respective model is retrieved.

However, aside from this slight automation of epochs, the parameters have been decided by trying different important parameters such as convolutional layers, filter number and so on. Examining the validation set the parameters have been selected with the minimum rMAE error.

First, the single camera model has the following filters. The dilated convolutional operator is connected to the inputs with a filter size of $k = 15$ dilated by $d = 4$ and step size of $s = 1$. This operator uses 128 filters to filter the initial image. The convolutional operator is set to a filter size of $k = 3$ and a step size of $s = 1$, and the pooling operator is set to window size of $k = 3$ and step of $s = 2$. Five convolutional layers are used following the described configuration with an ascending number of convolutional filters in the convolution operator, which starts at 32 and doubles each layer up to 512. The output of the five convolutional layers is connected to a fully connected layer with 512 neurons and ReLU activation. A dropout operator is set to the last fully connected layer and 75% of the neurons are disabled.

The multi-view model shares most of the hyper parameters described beforehand except for the configurations of the CNNs. The CNN_1 has three convolutional layers with filters ranging from 32 to 128, doubling at each layer and the CNN_2 add three more layers from 128 to 518. Kernel sizes, step, and so on, are the same as with the single camera model.

Baseline methods

To compare the results from CNN models, two baseline methods are evaluated along with the CNN proposals: the cloud fraction and random forest.

- **Cloud Fraction:** One of the most used approaches to estimation is the cloud fraction (CF) algorithm. As a baseline for comparison the cloud fraction has been computed, where irradiance is calculated as the result of the product of clear-sky irradiance and the coverage from the image (i.e. if there were no clouds, the irradiance would be maximum). Using a cloud detection algorithm (such as HYTA) the cloud coverage percentage is calculated. The cloud fraction approach is commonplace in physical irradiance estimation and yields powerful results in this context [Chen et al., 2012]. The CF forecasts have been provided by the MATRAS group.

- **Random Forest using Heinle features:** An approach to irradiance estimation through feature extraction is using the same technique presented on Section 3.1.2. Given that the features described earlier are suitable for cloud classification, it is hypothesized then that these same features will be able to estimate to some extent solar irradiance. The inputs to this random forest are features from 1 to 9 as detailed in Section 3.1.1, including all spectral features (average red and blue, blue standard deviation and skewness, average differences from RGB channels), textural features (blue energy, homogeneity, contrast and entropy) and coverage.

  A minor difference is that coverage here has been calculated using the HYTA algorithm result instead of the thresholding presented in the feature list. Given that the CNNs used include the cloudy pixels detected by HYTA, it would be unfair for the RF algorithm to use a coverage calculated through a less accurate method than HYTA.

  Given that there are three cameras available, and given that CNNs have been evaluated using a single image and multi-images, two models can be built: a single image model and a multi-view model both using RF

Table 3.7: Summary of metrics by cloud type for the single camera models

Metric		CF	RF	No Sun CNN	Single CNN
Global	rMAE	19.51	14.99	15.96	**14.29**
	rRMSE	27.22	21.03	20.8	**18.29**

for estimation. The first model, uses the random forest algorithm with the image features described as inputs to estimate irradiance. The second approach takes the features from all three images and concatenates them into a single vector. This vector contains the inputs to the RF algorithm.

3.2.4 Experimental Results

Here the results from the experimentation are presented and discussed. Several factors have been studied: the influence of using one vs multiple images, adding channels with the sun distance to pixels, and the improvement from CNNs over the baseline approach. A set of experiments have been set to study the effects of various design decisions of the CNN.

Experiments are divided in two categories, the single camera models and the multi-view models.

Single camera results

A summary of the results obtained by the single camera models can be seen in Table 3.7. This table shows the two explained metrics measured for each experiment carried out including the CF and the random forest model (RF), which are the baseline methods, along with a convolutional network without the sun distance channel (No Sun CNN) and the complete single camera model (Single CNN). The No Sun CNN has the purpose of testing whether the sun distance channel is relevant to improve regression. The nimbostratus cloud type represents an extreme outlier, with a representation on the test set of 1 instance. Given that a single example is not representative of a sky type it has not been included in the macro-average measure as it would be distorted. This comparisons allow examination of some of the properties of the Single CNN.

Examining the obtained errors some key differences between models can be seen. The first observation is the dominance of the Single CNN model over all the other alternatives evaluated. Both the baselines CF and RF underperform when compared to the Single CNN. The rMAE of RF is similar to the Single CNN, being the second best overall model. Concretely, for rMAE the improvement is of 5.22% (relative 26,77%) for CF and of 0.7% (relative of 4.66%) for RF. This pattern is repeated on rRMSE, however the improvement

Table 3.8: Analysis of rMAE metric for the single camera models by cloud type and each experiment. These values are expressed in %. The best results for each class and experiment are marked in bold. The final macro-average values are presented below the specific error metrics.

Sky Type	CF	RF	No Sun CNN	Single CNN
Clear Sky	**3.28**	8.57	8.6	6.4
Cirrus	8.77	7.2	8.3	**6.28**
Cirrostratus	**9.66**	13.28	12.99	11.06
Cirrocumulus	13.4	**6.85**	9.51	9.39
Altocumulus	26.87	16.75	20.15	**13.91**
Altostratus	28.99	8.42	12.48	**7.87**
Stratus	36.10	34.78	37.58	**30.63**
Stratocumulus	31.38	19.55	23.52	**19.07**
Cumulus	9.59	10.18	**7.44**	10.60
Multicloud	24.86	20.04	**18.59**	18.79
Aerosol	34.55	10.52	**9.78**	10.25
Macro-average	20.68	14.19	15.36	**13.11**

Table 3.9: Analysis of rRMSE metric for the single camera models by cloud type and experiment. The final macro-average values are presented below the specific error metrics.

Sky Type	CF	RF	No Sun CNN	Single CNN
Clear Sky	**3.62**	12.79	11.08	8.5
Cirrus	12.62	9.05	10.84	**8.32**
Cirrostratus	**12.32**	16.51	14.9	14.19
Cirrocumulus	15.48	**9.41**	11.87	11.20
Altocumulus	33.85	20.91	24.81	**17.32**
Altostratus	29.69	10.21	19.01	**9.77**
Stratus	45.96	44.37	45.51	**38.05**
Stratocumulus	41.11	26.22	28.5	**24.65**
Cumulus	14.35	12.28	**8.94**	12.94
Multicloud	31.12	29.06	**23.69**	25.89
Aerosol	35.99	**15.53**	16.43	16.58
Macro-average	25.1	18.76	19.6	**17.04**

of the Single CNN over RF is notably higher than before with an improvement of 2.74% (relative 13.03%). The improvement over the CF is higher than the improvement on RF, this indicates that the benefits of CNNs are higher on traditional statistical models than on other machine learning models.

Other notable observations lie in the performance of the No Sun CNN. This network is worse than RF and the Single CNN on rMAE, which shows the importance of the sun distance channel to improve performance of the Single CNN. An improvement of 1.67% (relative 10.46%) is achieved on rMAE and 2.51% (relative 12.61%) on rRMSE.

In Tables 3.8 and 3.9 the rMAE and rRMSE results are broken down by sky condition (i.e. cloud type) and the macro-average is shown. As seen in Table 3.6, various cloud types differ in average GHI and K_t, therefore this analysis reveals how complex it is for the model to estimate the respective irradiance for each of the sky types. The nimbostratus cloud type has been removed due to the low representation in test, as mentioned before.

Several remarks can be drawn from the rMAE table. The first is that the Single CNN outperforms in terms of macro-average every other proposal (see the last row of Table 3.8) by a notable margin, with an improvement

of 7.57% (relative 36%) on CF and of 1.08% (relative 7.61%) for RF. The Single CNN keeps outperforming both, the RF and CF baselines and the No Sun CNN. Including sun distance improves the macro-average rMAE of the CNN, which reinforces the notion that the sun distance channel contains relevant information about irradiance. These observations are similar when the macro-average rRMSE is evaluated.

Analysing errors of different methods on specific sky types (Tables 3.8 and 3.9), the following results are observed. The initial observation is the dominance of the Single CNN model over the CF baseline. Except for clear skies and cirrostratus, the Single CNN model outperforms CF in every other cloud type. A similar phenomenon happens when RF is compared to the Single CNN, where, except for cirrocumulus, the Single CNN performs better on most other types. There are some cloud types where performance is very similar, such as aerosols or cumulus where depending on the metric (rMAE or rRMSE) the first is better than the latter and vice versa. It is also worth noting that the Single CNN gives the most accurate estimations on 7 out of the 11 sky types shown compared to the baselines. These observations are mostly the same on rMAE and rRMSE.

When the No Sun CNN and the Single CNN are evaluated together further observations can be drawn. The results here prove to favour the Single CNN, while the No Sun CNN is better at cumulus, multiclouds and aerosols, the differences are slight and the rest of sky types are still dominated by the Single CNN.

There are some additional interesting observations worth analysing. The CF model shows very robust results on clear skies, outperforming drastically all other models evaluated. The stratus class is very hard to estimate for most models, with very high error. The Single CNN manages to lower this error, but remains relatively high compared to any other cloud type. The global metrics are higher than the macro-average errors across all experiments, this may happen due to overrepresented sky types being difficult examples to estimate.

Multi-view results

A summary of the results obtained by the Multi-view models can be seen in Table 3.10. The baselines (CF and RF) presented before conform two of the experiments, except that now RF also use three cameras. The rest of the experiments are further alternatives of CNNs: the average single camera model (Single Avg CNN), which just averages the three views, the best performing CNN (Multi Max CNN) but removing the sun distance channel (No Sun CNN), and the proposed architecture with an average or max combina-

Table 3.10: Summary of metrics, total and averaged by cloud class for the multi-view models. The average of rRMSE and rMAE of cloud types excludes nimbostratus.

Metric		CF	RF	Single Avg CNN	No Sun CNN	Multi Avg CNN	Multi Max CNN
Global	rMAE	19.51	14.14	14.78	14.07	14.38	**13.08**
	rRMSE	27.22	19.55	19.72	18.85	19.44	**17.86**

Table 3.11: Analysis of rMAE metric for the multi-view models by cloud type and for each experiment. The best results for each class and experiment are marked in bold. The final macro-average values are presented below the specific error metrics.

Experiment	CF	RF	Single Avg CNN	No Sun CNN	Multi Avg CNN	Multi Max CNN
Clear Sky	**3.28**	6.4	8.18	12.75	7.86	8.21
Cirrus	8.77	**6.28**	11.14	8.89	10.72	8.98
Cirrostratus	9.66	11.06	11.87	**8.74**	9.05	8.94
Cirrocumulus	13.4	9.39	**4.98**	9.1	12.06	8.7
Altocumulus	26.87	13.91	13.86	**13.69**	16.31	14.26
Altostratus	28.99	**7.87**	13.74	11.81	12.55	11.14
Stratus	36.1	30.63	26.91	33.13	**22.35**	23.02
Stratocumulus	31.38	19.07	19.86	19.25	21.31	**17.35**
Cumulus	9.59	10.6	9.75	6.05	**4.93**	8.87
Multicloud	24.86	18.79	18.06	18.12	17.07	**15.98**
Aerosol	34.55	10.25	8.34	11.36	10.24	**6**
Macro-average	20.68	13.11	13.34	13.9	13.13	**11.95**

tion (Multi Avg CNN and Multi Max CNN). The impact of the combination operator of the three views is examined by comparing the max and average operations. As with the single camera model, the distance to the sun is also examined (No Sun CNN = Multi Max - Sun distance channel). The capabilities of the single camera model that are examined through the average model as explained in Section 3.2.1. This architecture consists in averaging the convolved image outputs of CNN before the fully connected layer.

New patterns are found in the multi-view table for rMAE. First, the Multi Max CNN outperforms every other alternative tested. The benefits over baselines is clear, with an improvement of 6.43% (relative 32.96%) for CF and 1.06% (relative 7.5%) for RF. The rest of CNNs perform with similar efficacy as the RF model, with improvements in the range from 0.99% to 1.7% (relative 7% to 11.5%). The Multi Max CNN also outperforms the rest of the models with similar absolute improvements when rRMSE is considered.

The Single Avg CNN is worse than the RF baseline on both rRMSE and rMAE, which supports the idea that the combination of intermediate features has to take place between convolutions. The Multi Avg CNN has an average performance, being better than RF on rRMSE but worse on rMAE. It is also concluded from these results that the maximum intermediate combination is preferable to the average combination. The distance from the sun also improves performance by a visible margin.

In Tables 3.11 and 3.12 the rMAE and rRMSE results are broken down by sky condition. Again as with the single camera results the nimbostratus class has been excluded.

Table 3.12: Analysis of rRMSE metric for the multi-view models by cloud type and each experiment. The final macro-average values are presented below the specific error metrics.

Experiment	CF	RF	Single Avg CNN	No Sun CNN	Multi Avg CNN	Multi Max CNN
Clear Sky	**3.62**	8.5	12.05	18.78	11.53	13.62
Cirrus	12.62	**8.32**	14.01	12.1	13.65	11.85
Cirrostratus	12.32	14.19	16.06	**10.97**	11.67	11.05
Cirrocumulus	15.48	11.2	**6.5**	10.03	16.63	10.27
Altocumulus	33.85	17.32	18.24	18.01	20.25	**16**
Altostratus	29.69	**9.77**	16.52	15.32	17.84	12.99
Stratus	45.96	38.05	36.68	42.75	31.21	**30.08**
Stratocumulus	41.11	24.65	25.31	24.37	25.77	**23.23**
Cumulus	14.35	12.94	11.49	7.68	**6.21**	11.35
Multicloud	31.12	25.89	23.17	23.12	22.24	**21.5**
Aerosol	35.99	16.58	12.57	13.32	13.1	**11.44**
Macro-average	25.1	17.04	17.51	17.86	17.28	**15.76**

The results show interesting patterns for rMAE metrics. Again, the Multi Max CNN outperforms the rest of tested methods when the macro-average metrics are considered (last row of Tables 3.11 and 3.12). Considerable improvements are seen with respect to the baselines, with an error reduction of 8.73% (relative 42.21%) for CF and of 1.16% (relative 0.88%). The Multi Max CNN also improves performance over the rest of CNNs, with improvements ranging from 1.1% to 2% (relative 8% to 14%). With this, it is again concluded that the inclusion of the sun distance is very relevant to estimation, that the combination of convolutional outputs in the middle of the architecture detects important features within the image and, finally that the optimal way to combine these outputs is the maximum operation.

The specific sky type error metrics also show interesting observations. When compared to the baselines, the Multi Max CNN model visibly outperform the CF and RF on most sky types. The CF is the best at estimating clear skies, but it is fairly poor on all other models, being worse in every scenario than the Multi Max CNN. The RF model suffers in a similar way, only being better at stimating cirrus and altostratus. Any other type is better estimated by the Multi Max CNN. On the other hand, the other CNNs have some strengths, for example, the Single Avg CNN is the best at estimating cirrocumulus. However, this happens with two cloud types at a maximum, while the Multi Max CNN is the best alternative on most cloud types.

On rRMSE (Table 3.12) the observations from rMAE remain mostly the same, however, the comparison of CNNs is slightly different. For rRMSE the CNNs excel at only one cloud type, and the Multi Max CNN is better on the rest of sky types.

Interestingly enough the CF model remains the best alternative for clear sky estimation. The error metrics on the Multi Max CNN usually range from 15% to 10% except for some difficult cloud types such as stratus, stratocumulus and multiclouds.

Table 3.13: Comparative of rMAE and rRMSE error metrics of the single camera model and multi view model. The final macro-average and global values are presented below the specific error metrics.

Metric	rMAE		rRMSE	
Experiment	Single CNN	Multi Max CNN	Single CNN	Multi Max CNN
Clear Sky	**6.4**	8.21	**8.5**	13.62
Cirrus	**6.28**	8.98	**8.32**	11.85
Cirrostratus	11.06	**8.94**	14.19	**11.05**
Cirrocumulus	9.39	**8.7**	11.20	**10.27**
Altocumulus	**13.91**	14.26	17.32	**16**
Altostratus	**7.87**	11.14	**9.77**	12.99
Stratus	30.63	**23.02**	38.05	**30.08**
Stratocumulus	19.07	**17.35**	24.65	**23.23**
Cumulus	10.60	**8.87**	12.94	**11.35**
Multicloud	18.79	**15.98**	25.89	**21.5**
Aerosol	10.25	**6**	16.58	**11.44**
Macro-average	13.11	**11.95**	17.04	**15.76**
Global	14.29	**13.08**	18.29	**17.86**

Multi-view and Single Camera Comparison

The results between the final proposal for single camera estimation (Single CNN) can be compared with the final multi-view model (Multi Max CNN). Table 3.13 contains a comparative of all the results previously shown for the single camera and multi-view models. Both the rMAE and rRMSE are analysed for both models.

The improvements on global and macro average metrics can be observed for rMAE and rRMSE. The multi-view model achieves an improvement of 1.16% (relative 8.84%) on macro-average rMAE, and an improvement of 1.21% (relative 8.47%) on global rMAE. On the other hand, improvements are slightly lower on rRMSE with a 1.28% (relative 7.75%) on macro average and 0.43% (relative 2.35%). The higher improvement on macro-average may represent the superior generalization capabilities of the multi-view estimation.

Further comparison can be drawn between specific sky types, where the multi Max CNN is always the best on most cloud types except clear skies, cirrus and altostratus. There is no best option for the altocumulus type, as the rMAE estimation error of the Single CNN is better, but the multi-view model performs best when rRMSE is measured. Interestingly enough, the standard deviation of the metric values of the multi-view models is lower (5.11 for rMAE and 6.7 for rRMSE) than the deviation of the single camera models (7.79 for rMAE and 9.7 for rRMSE), which may indicate that the multi-view models are more consistent across cloud types.

A qualitative comparison is shown on Figure 3.12 where two scatterplots are shown (predicted vs. actual GHI). The first scatterplot represents the Single CNN and the second the Multi Max CNN. They both fit the diagonal with relative accuracy, however the multi-view model fit is somewhat tighter.

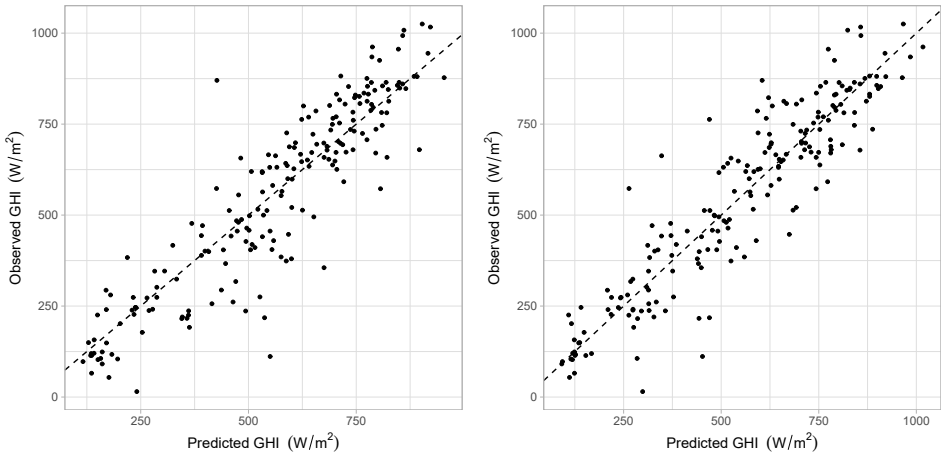

Figure 3.12: Scatterplots of the proposed CNN models. Left plot: Singe image model scatterplot. Right plot: Multi-view model scatterplot

3.2.5 Summary of experimental conclusions

A CNN approach to irradiance estimation is proposed in this section. It is shown that the proposed CNN outperforms other methods and architectures in most scenarios tested for the single camera and multi-view approaches.

The proposed CNN architecture has been compared with two baseline methods (Cloud Fraction and Random Forest). Also, a study of how the different elements of the proposed CNN impact results, has also been carried out. The final proposal explained in Section 3.2.1 performs better than any other tested approach. While improvements are lower on rRMSE than on rMAE, the final CNN architecture is still better.

The inclusion of the additional information channels (sun distance and cloudy pixels) have shown positive results and it has been tested that the removal of sun distance negatively impacts metrics. Also, results are also worse when convolutional layers are removed and when a different combination of the three points of view is used (in the case of the multi-view architecture). Thus, all elements present in the proposed architecture prove to be important for the final result. This happens on both approaches, the single camera and the multi-view model. Importantly, it is additionally shown that the combination of multiple image perspectives improves the accuracy of the machine learning model.

As with cloud classification, there are some differences when macro-average metrics are used. In this case macro-average metrics are lower than the global metrics. This means that some over-represented cloud types are badly estimated. Additionally, improvements on macro-average metrics are slightly higher that global metrics.

To sum up, the proposed CNN shows an improvement over traditional physical techniques (Cloud Fraction) and other feature-based machine learning methods (Random forest with feature extraction), specially when multiple images are combined.

Chapter 4

Integration of irradiance predictors for short-term forecasting

Solar forecasting is crucial for the energy grid [Kaur et al., 2016]. There are many physical models, ranging from simple approximations such as persistence, a simple but accurate approach for very short term forecasts, to more sophisticated Numerical Weather Predictions (NWP) such as WRF-Solar, which runs a simulation of atmosphere dynamics to forecast irradiance. Machine learning has been used to improve irradiance forecasts with varying degrees of success. Forecasting models predict irradiance specialize at certain prediction horizons. For example, Persistence only offers good results on short term predictions and WRF-Solar has peak performance on far prediction horizons [Jimenez et al., 2015]. Therefore, a model can be developed to combine a set of diverse forecasts to get a more accurate final model. This technique is called forecast integration (or blending) and consists on taking narrow already-available predictions and optimally combining them into a single value.

Although this issue is not new in the framework of weather forecasting research [Vislocky and Fritsch, 1995], very few works have conducted this kind of analysis and, even less for short-term solar irradiance forecasting. Additionally, models of different nature, that perform better at diverse forecasting horizons, have seldom been combined in the literature.

Therefore, in this thesis a machine learning model is developed for the integration of GHI and DNI of four predictor models across four meteorological stations within the region of the southern Iberian peninsula. The importance of forecasting horizons in the development of integration models is discussed, where different approaches are identified. A Horizon approach, building a machine learning model for each prediction horizon is proposed along with a General approach, building a single machine learning model for all horizons simultaneously.

For system operators, forecast restricted to a single location may not be enough for the optimal management of the energy grid. Therefore, forecasts are needed at regional scopes [Renné, 2014, Pierro et al., 2017]. This kind of forecast is important, as in some situations, only the aggregated irradiance over a region is of interest for practical applications.

In this thesis, two regional approaches for GHI and DNI blending are identified, an average of the integrated predictions from the individual stations and a machine learning model combining all predictor models from the four stations.

The forecast integration models are described in this chapter as follows. First the blending approaches are described in Section 4.1 presenting the Horizon, General and regional aggregation models. There are four predictor models used by the integration algorithms that are described in further detail in Section 4.2. The machine learning models are built using Support Vector Regression (SVR) as described in Section 4.3. How the data was gathered is described in Section 4.4, followed by the experimental methodology in Section 4.5. The results of this work are shown in Section 4.6 and discussed in Section 4.7.

4.1 Description of blending approaches

The concept of blending and integration is briefly introduced in Chapter 2 Section 2.3.3. Forecast integration is an approach to forecasting, that uses similar concepts from time series forecasting, such as current time t or horizon h. In broad terms integration takes the results of other prediction models and, using statistical transformations, builds a new prediction by combining predictions.

The main difference with time series forecasting is that, instead of finding a function that combines a set of historical values ($V(t-1), V(t-2), ..., V(t-p)$), a set of predictions are used instead of measures ($M_1(t,h), M_2(t,h), ..., M_n(t,h)$). Each prediction model M_k can give a forecast at a time t for a horizon h, therefore the first model M_1 in (t,h) makes a prediction at time t looking ahead into the future h minutes. A simple forecast integration approach is shown in Equation 4.1 where four models are used to forecast.

$$P(t+h) = f(M_1(t,h), M_2(t,h), M_3(t,h), M_4(t,h)) \tag{4.1}$$

A function f can be built that combines all models. Here, instead of a linear combination of forecasts, as described in the literature, the combination

function is found via machine learning. In this case the algorithm used is SVR, as described in Section 4.3.

For example, if at minute $t = 0$ the irradiance at $h = 30$ minutes ahead is to be forecasted ($P(0 + 30)$), each model would give its respective value ($M_k(0, 30)$) and would be combined by the function f to return the final forecast.

The predictions of the four models are described in Section 4.2. Horizons have a great impact on traditional forecasting so it is assumed that there will be an important impact on prediction integration and a way to divide horizon data is decided on this assumption.

The blending approach aims to calculate at a given point in time t an accurate prediction of irradiance (either GHI or DNI) for different forecasting horizons (h). Two different approaches have been used to optimize the combination of the four predictors. The main difference between them is that the Horizon approach constructs a model for each horizon, while the General approach trains a unique model valid for all horizons. Both are described below.

4.1.1 Horizon approach

The Horizon approach consists in training a machine learning model f_h for each horizon h. This approach allows horizon-specific synergies to be detected by the learning algorithm, building specialized models for different horizons. The horizon forecasting assumes that there are inherent differences in a set of predictions for different forecasting horizons. This is the common and straightforward approach to traditional forecasting.

There is a model for horizon 15, another for horizon 30, and so on, up to horizon 360, which results in 24 different models. Figure 4.1 displays how each model f_h is trained using data belonging to horizon h only. Thus, input data for each model f_h is a tuple with shape $((M_1(t, h), M_2(t, h), M_3(t, h), M_4(t, h)), P(t + h))$ for all available time points. Here M_k denote predictors available at time t, that make forecasts at time point $t + h$ with a prediction horizon of h. Once the models have been constructed, Equation 4.2 shows how the 24 f_h models can be used for irradiance forecasting. For instance, if at time t it is desired to know the forecast in 15 minutes time, f_{15} would be used and $f_{15}(M_1(t, 15), M_2(t, 15), M_3(t, 15), M_4(t, 15))$ would be computed.

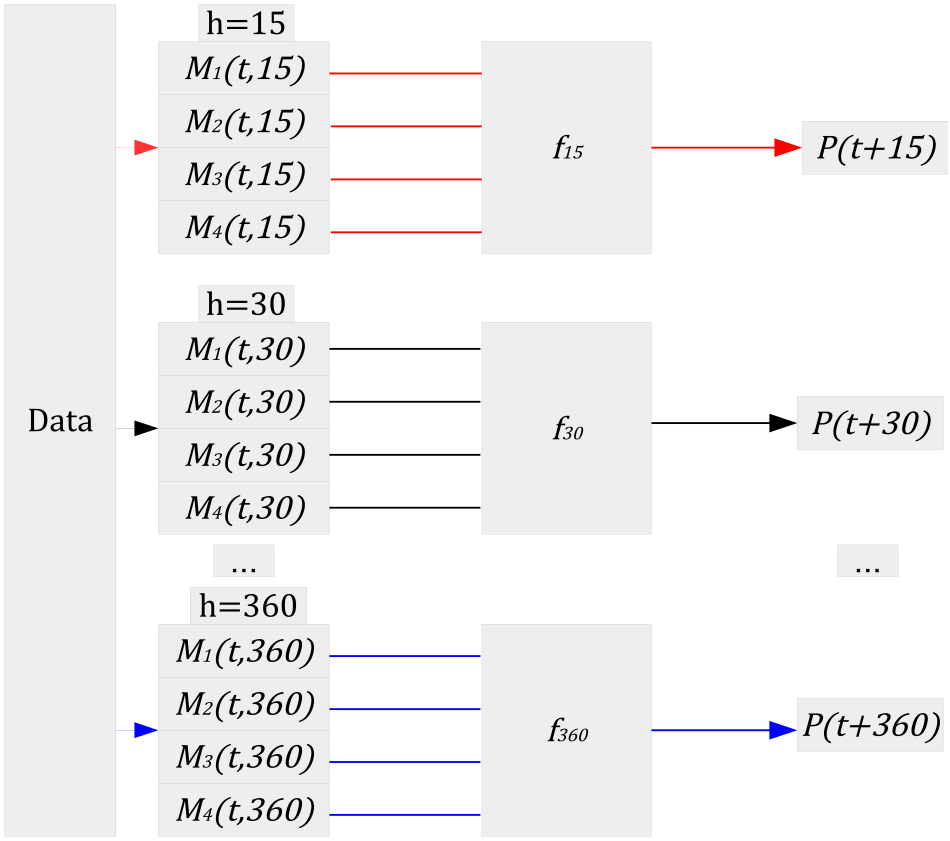

Figure 4.1: Visual representation of data flow and prediction making process for the Horizon approach.

$$P(t+h) = \begin{cases} f_{15}(M_1(t,15),...,M_4(t,15)), & \text{if } h = 15 \\ f_{30}(M_1(t,30),...,M_4(t,30)), & \text{if } h = 30 \\ \dots & \\ f_{360}(M_1(t,360),...,M_4(t,360)), & \text{if } h = 360 \end{cases} \tag{4.2}$$

4.1.2 General approach

The General approach takes a different direction compared to the Horizon approach by constructing a single model for all horizons simultaneously. This approach allows the learning algorithm to detect inter-horizon relationships by building a model based on the complete dataset, joining data from all horizons at a station.

The model built with this approach will forecast irradiance independently of the prediction horizon, so a single function is needed to operate. This principle is represented in Figure 4.2 where it is shown how data from all horizons is combined and a single model f is trained. In this case, the training data will consist of tuples with shape $((M_1(t,h), M_2(t,h), M_3(t,h), M_4(t,h)), P(t+h))$, for all available time points and horizons. Once the model f is trained, it is used for irradiance prediction using the Equation 4.3. It is noted that, although $M_k(t,h)$ is used as input for predictions at $t+h$, f itself is independent of h. Compare with Equation 4.2.

$$P(t+h) = f(M_1(t,h), M_2(t,h), M_3(t,h), M_4(t,h)) \tag{4.3}$$

4.1.3 Regional blending

The aim of regional blending is to predict the average of irradiance (GHI and DNI) over a region, which in this case, it will be the average of the measured irradiances of the four local stations. Seville, Jaen, Lisbon and Madrid, give a fair representation of the southern half of the Iberian peninsula region.

Two different regional blending approaches are studied to predict irradiance at the different forecasting horizons considered. The first one performs regional forecasting by computing the average the four station blending models. And the second one builds a model by using all sixteen inputs in the dataset as predictors (that is, four physical models times four locations).

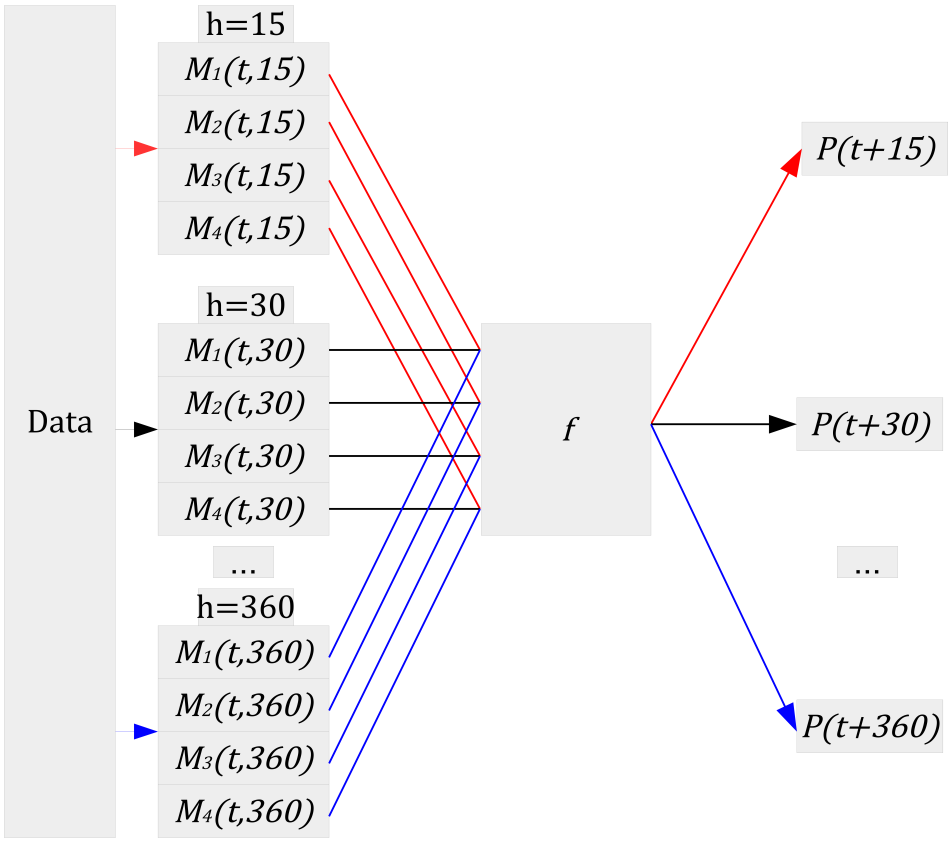

Figure 4.2: Visual representation of data flow and prediction making process for the General approach.

Average approach for regional blending

The average regional approach is a statistical average of the models of the four stations. It averages the predictions at a time point to cancel out the errors made by the individual local predictions. This technique is known to produce acceptable forecasts on a regional scale on forecasting.

This principle can be seen in Equation 4.4 where the machine learning predictions P^s from four stations are averaged for each $t + h$ available. The first row in Equation 4.4 represents the General approach of Equation 4.3, for each of the four stations (where s = Seville, Jaen, Lisbon, or Madrid). The Horizon approach expression could have been used in its place (Equation 4.2) to predict on a regional scale with either local techniques.

$$\begin{aligned} P^s(t+h) &= f(M_1^s(t,h), M_2^s(t,h), M_3^s(t,h), M_4^s(t,h)) \\ P^R(t+h) &= \frac{1}{r}\sum_{s=1}^{r} P^s(t+h) \end{aligned} \tag{4.4}$$

where s represents the station for which the prediction P is made. Here, r represents the number of stations being processed, in this case $r = 4$. Finally the regional prediction is represented as P^R.

Integration approach for regional blending

The integrated regional model constructs a machine learning model whose inputs are the four available predictors at each of the r stations. Therefore, r stations x 4 predictors will be used as inputs. The target output to train model f is now directly the average of irradiance (GHI or DNI) at the four locations.

This integration approach is represented in Equation 4.5 where a single model is generated for all sixteen predictions. As with the average approach to regional integration, this representation resembles closely the General approach expression. Nonetheless, the Horizon approach could also have been followed by using a different f_h for each horizon h.

$$\begin{aligned} P^R(t+h) = \; & f(M_1^1(t,h), M_2^1(t,h), ..., M_4^1(t,h), \\ & M_1^2(t,h), M_2^2(t,h), ..., M_4^2(t,h), \\ & \ldots, \\ & M_1^r(t,h), M_2^r(t,h), ..., M_4^r(t,h)) \end{aligned} \tag{4.5}$$

where M_k^s is the forecasting model k at station s.

4.2 Description of predictors

Four forecasting models have been used as predictors: Smart Persistence, Satellite, WRF-Solar and CIADCast. All models described here are configured to work up to six hour ahead, 15 minutes time resolutions and to produce GHI and DNI forecasts. Predictor models are are based on diverse foundations, each reaching best performance in different time-frames.

The operative parameters are described. All the data extracted from these models has been provided by the MATRAS group.

4.2.1 Persistence

Persistence is calculated by assuming the current measured irradiance I_0 at a time t remains constant, therefore a predicted irradiance I at a time point $t+h$ is the same at time t. This is seen in Equation 4.6.

$$I(t+h) = I_0(t) \tag{4.6}$$

Persistence can be improved by using the clear sky ration between measured irradiance and clear-sky irradiance, this approach is known as Smart Persistence. This model is computed with the actual measured irradiance I_0 and corrected with the variation of the clear-sky (cs) irradiance I_{cs} from the initial time t to a future time $t+h$. The relation between current irradiance and cs irradiance at a certain time t is kept constant and multiplied by the clear-sky irradiance in future $t+h$. The European Solar Radiation Atlas (ESRA) [Rigollier et al., 2000] cs model is used (Equation 4.7).

$$I(t+h) = \frac{I_0(t)}{I_{cs}(t)} \cdot I_{cs}(t+h) \tag{4.7}$$

Smart persistence performs best at short-term horizons, usually being the best forecasting model at 15 min horizon.

Satellite Model

In this method, satellite images are first processed to derive the so-called cloud index images, an intermediate step to retrieve the clear-sky index images and then the solar irradiance maps [Rigollier et al., 2004]. Secondly, a statistical comparison of various consecutive cloud index (CI) images allows deriving the cloud motion vector field. The open source Particle Image Velocimetry (OpenPIV) [Mori and Chang, 2003] algorithm, with a square window of 40 km, has been used to estimate the cloud motion vectors (CMVs). This

methodology to derive the CMVs was found to be the optimal for the study region [Rodríguez-Benítez et al., 2016].

The discrete cloud motion vector field is transformed into a continuous flow computing the streamlines, i.e., a family of curves tangent to this wind field [Nonnenmacher and Coimbra, 2014]. The streamline passing through the station location is used to obtain the future cloud index values, then the clear-sky index values and, finally, the GHI forecast.

Satellite models are powerful on short-term horizons, usually better than smart persistence after the first 15 to 30 minutes.

4.2.2 Weather Research and Forecast Model (WRF)

NWPs use mathematical models based on physical principles of the atmosphere and oceans to predict the weather based on current weather conditions. WRF-Solar [Jimenez et al., 2015] is a particular physical configuration of the WRF numerical weather prediction model version 3.6 devised for solar energy applications. It has the improved parameterizations from the Thompson aerosol-aware microphysiscs scheme [Thompson and Eidhammer, 2014] for the interactions of solar irradiance with clouds and aerosols.

Further improvements are included with the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) short- and long-wave irradiance parametrizations [Iacono et al., 2008], which are coupled with cloud physics parametrization. Direct effect of rural aerosol is taken into account in order to fully couple the cloud-aerosol-radiation system. Other parametrizations used were the revised MM5 Monin-Obukov surface layer parametrization [Jiménez et al., 2012], 5-layer thermal diffusion and the Mellor-Yamada Nakanishi Niino Level 2.5 planetary boundary layer parametrization [Nakanishi and Niino, 2009]. A new option for shallow convection [Deng et al., 2014] is included in WRF-Solar, thus the cumulus parametrization was switched off.

To extract the forecasts, the model was configured with one domain of 5 km spatial resolution and 50 vertical levels Initial and boundary conditions were taken from the National Centers for Environmental Prediction (NCEP) GFS [0.5 Deg] data set (gov.noaa.ncdc:C00634). For each day of the study data set, three 18-h simulations were run starting at 00, 06 and 12 UTC, discarding the first 6 simulated hours as spin-up. Outputs were saved every 15 minutes, at the moment when the satellite passed over each evaluation station.

The WRF-Solar due to its complex simulations underperforms at short-term horizons, but is very powerful in long-term horizons after 3 hours.

4.2.3 CIADCast Model

The CIADCast model [Arbizu-Barrena et al., 2017] for short-term solar irradiance forecasting is based on the advection and diffusion of cloud index estimates derived from satellite using the WRF [William et al., 2008] NWP models. Cloud index values are inserted in the WRF cell which corresponds to the cloud top height provided by the EUMETSAT product. Then, WRF is used to advect and diffuse the cloud index values as dynamical tracers both horizontally and vertically.

CIADCast forecasts were derived from the WRF-solar runs explained in section 4.2.2. In short, CI retrievals from satellite are advected and diffused using the WRF model to obtain cloudiness forecasts and therefore irradiance forecasts.

After the model is run, the sum of each column of cloud index values is computed to obtain again a two-dimensional cloud index map. The CI forecasts values at the stations locations are transformed to GHI and DNI as with the satellite model.

For each day, three 18-h simulations were run starting at 00, 06 and 12 UTC, discarding the first 6 simulated hours as spin-up. The model is stopped at the satellite retrieval time, then the CI image is inserted, and restarted to simulate the next 6 h. Outputs were saved every 15 min.

CIADCast performs best on far horizons, performing better as horizons increase, usually after 3 hours.

4.3 Support Vector Regression

Support Vector Machines (SVM) [Drucker et al., 1997] is a machine learning algorithm often used for classification and regression problems. For classification, the learning algorithm searches the optimal hyperplane dividing two different classes. This is achieved maximizing the margin between the instances of both classes.

Using what is called a kernel function, the data is transformed into a higher dimensional space where data can be separated as usual using a hyperplane. In fact, kernels do this transformation implicitly, which allows to use infinite-dimensional spaces, as those implicitly used by gaussian kernels. Non-linear kernels allow for non-linear boundaries.

The idea behind SVM can be used for multi-class classification or regression. The regression algorithm as explained by [Smola and Schölkopf, 2004] is the following. Let‘s assume a training data set $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

containing real values, y_i being the desired outputs and x_i the input variables. The goal is to find a function $f(x)$ that has less than a ε deviation from the observed y_i values. The equation $f(x) = \langle w, x \rangle + b$ is the general form of the linear model sought, where $\langle \cdot, \cdot \rangle$ is the dot product, b a real number and w is the weights vector, to be optimized. The objective then is to look for a small w, and this can be done by minimizing $\langle w, w \rangle$ by means of the optimization problem represented in Equation 4.8.

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} \cdot \|w\|^2 \\ &\text{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \end{aligned} \tag{4.8}$$

However, Equation 4.9 assumes the existence of a valid solution that approximates all (x_i, y_i) pairs (within an allowed error ε). This may not always be the case and some minor errors should be allowed in order to find a feasible solution. This is solved introducing slack variables ξ_i and ξ_i^* (errors beyond allowed error). The final optimization problem is presented in Equation 4.9:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} \cdot \|w\|^2 + C \cdot \sum_{i=1}^{n} (\xi_i + \xi_i^*) \\ &\text{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \tag{4.9}$$

where $C > 0$ is a constant that determines by how much the ξ_i and ξ_i* should be taken into account into the optimization process. This optimization problem can be further extended to non-linear models by means of the kernel trick. One of the most widely used kernels is the gaussian one. In that case, the SVM model becomes that of Equation 4.10:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \cdot e^{-\frac{\|x - x_i\|}{2\sigma^2}} \tag{4.10}$$

The SVR algorithm is versatile, powerful and efficient, as it can find linear and non-linear relationships in large datasets faster than most ANNs and Ensemble models. The nature of the solar forecasting problem is, so far, unknown. Therefore, both linear and-non linear approaches have to be tested. SVR is sensitive to extreme outliers and the magnitude of data.

Figure 4.3: Digital elevation map of the Iberian peninsula, dots represent the studied stations.

4.4 Data set generation

Data collected at four radiometric stations representatives of the central (Madrid station), southern/southwestern (Seville and Jaen) and western (Lisbon) areas of the southern region of the Iberian Peninsula are included, the locations represented in Figure 4.3.

The atmospheric circulation over the study area is ruled by the semi-permanent Azores Islands high pressure center [Castro-Díez et al., 2002, Trigo et al., 2002, Trigo et al., 2004]. As a consequence, the study area shows a marked seasonal climate variability, associated with changes in the location and intensity of the Azores high. This marked seasonality is clearly observed in the solar irradiance variability [Pozo-Vázquez et al., 2004, Pozo-Vazquez et al., 2011, Santos-Alamillos et al., 2012]. In a recent work [Rodríguez-Benítez et al., 2018] the weather pattern associated to the different solar irradiance modes of variability in the study region were analysed.

From a topographic point of view, the study area may be split into two different parts. The western area is a nearly homogeneous at region, with wide valleys open to the Atlantic Ocean. The Lisbon and Seville stations are located in this part. On the other hand, the central and southern area shows a more complex topography, with several mountain ranges in the south (where the Jaen station is located) and a plateau in the center (where the Madrid station is located). In addition, there are some local weather features, such as cloudiness associated to the availability of moisture in coastal areas (Lisbon) or the presence of orographic clouds in areas close to mountain ranges. Nevertheless, synoptic weather patterns mostly dominate the cloudiness uniformly over the entire study area [Rodríguez-Benítez et al., 2018].

Table 4.1: Dataset example for Seville data

Station	Date	Hour	Horizon	SmartPers	Satellite	CIADCast	WRF-Solar	Measure
Seville	2015-03-03	10:15	15	299.72	620.9	230.3	226.29	283.1
Seville	2015-03-03	10:15	30	427.23	649.3	240.42	283.04	254.55
Seville	2015-03-03	10:15	45	627.45	674.16	249.23	326.29	303.18
.....	.....	....	.....	.....	.....	.....	.....	.....
Seville	2015-03-03	10:15	360	71.80	312.25	119.01	31.23	280.22
Seville	2015-03-03	10:30	15	417.26	649.01	295.54	296.93	254.55
Seville	2015-03-03	10:30	30	666.79	673.98	306.37	401.20	303.18
Seville	2015-03-03	10:30	45	636.63	619.67	445.28	599.42	347.69
.....	.....	....	.....	.....	.....	.....	.....	.....
Seville	2015-03-03	10:30	360	471.29	556.22	298.87	411.25	546.71
.....	.....	....	.....	.....	.....	.....	.....	.....

4.4.1 Data set description

The raw data has certain imperfections and some preprocessing has been required for the proper functioning of the proposals. The following measures are taken to construct the irradiance forecast blending dataset.

One minute time resolution solar irradiance data collected at the four ground stations were used. Both measurements GHI and DNI are available at all four stations. The observations cover the period March 2015 to February 2017, inclusive, (i.e., two years). Radiometers description and data quality procedure are fully detailed in [Rodríguez-Benítez et al., 2018]. Data associated with solar zenith angle (SZA) > 75 were discarded. Values corresponding to the images retrieval times at each station locations were used.

Forecasts of GHI and DNI up to 6 hours ahead, with a time step of 15 minutes, are obtained based on the four different models described beforehand, including: Smart Persistence, Satellite-based, CIADCast, and WRF-Solar. There are four different numerical inputs (four predictors) and a target (measure column).

The number of records from the raw source is different for WRF-Solar, CIADCast and Satellite. Smart Persistance records can always be retrieved if there is a real measurement but the other three methods may produce missing values. To produce a consistent dataset, when there is a missing prediction at time t for horizon h the complete record is discarded. In other words, the predictions of the other three methods are also discarded, even if they produced a forecast. The resulting dataset has always four predictors available, as there is no missing data. An example of the structure of the final dataset is shown in Table 4.1. This structure is repeated for all four stations.

Table 4.2 shows the number of forecasts available (for each model) as function of the forecasting horizon. A total of 116458 forecasts were obtained, ranging from about 7800 for the first horizons to less than 1500 at the 6 hours forecasting horizon. Note that there is significantly less data as the horizon increases, due to the time range and the increasing time window from t to

Table 4.2: Forecasts obtained at each forecasting horizon

Horizon (min)	N	Horizon (min)	N
15	7858	195	4664
30	7773	210	4256
45	7654	225	3904
60	7515	240	3558
75	7329	255	3262
90	7108	270	3009
105	6836	285	2764
120	6528	300	2492
135	6230	315	2218
150	5843	330	1956
165	5454	345	1716
180	5044	360	1487
Total	116458		

$t + h$.

There are no extreme outliers in the proposed dataset, however the magnitude of data has to be normalized in pre-processing between 0 and 1 for the SVR algorithm to function properly.

Here, instead of forecasting raw irradiance, the K_t is predicted. Forecasting K_t is easier than predicting the raw measurement of irradiance. The ratio K_t is the proportion of measured irradiance divided by clear sky irradiance. This measure removes the variations resulting from the daily cycle, which means the model becomes simpler (the daily cycle needs not be inferred from the data).

4.5 Experimental methodology

The methodology followed for the experimentation requires several decisions to be made: The hyper-parameter tuning for SVR (only for radial kernel) algorithm, how the division in train and testing data is performed, and which error metrics are used to compare the different models.

4.5.1 Cross Validation

Every approach described in Section 4.1 require models to be trained and validated. In this work, Cross-Validation (CV) has been applied for this purpose. CV is a common practice in machine learning validation. However, some modifications have been introduced that differ from the standard CV, which improve the fairness of the validation. CV creates K random partitions (or folds) of the same data size. For each partition z, a model is trained with all partitions but z, and tested on z. The CV result is the average of the K testing partitions.

Standard CV partitions the dataset randomly. This is appropriate when instances are independent but in the case of forecast integration, there is a

temporal dependency between instances. Standard CV risks using consecutive days for both train and test, which may result in overly optimistic estimations of performance. The nature of solar energy forecasting and the used dataset makes it very sensitive to this kind of issue.

In order to mitigate this problem, a variation of CV has been used here (Grouped CV). There will be 4 different partitions (CV with $K = 4$), one for each week of every month. Therefore, partition 1 contains the first week of January, the first of February, and so on. Similarly for partition 2, with the second week of every month. This CV variation guarantees that, at least, training and testing partitions will never contain instances belonging to the same week, which is enough for the purpose of a fair validation.

4.5.2 Hyper-parameters

Blending approaches can use both linear and non-linear models. For this purpose, experimental results have been obtained using linear and radial SVR. This learning algorithm requires the optimization of the hyper-parameter C (see Equation 4.9). A brute force search for the optimal C value has been conducted by testing a wide enough range of values. This has been done evaluating different C values ($C = (0.25, 0.5, 1, 2, 4)$) by performing standard CV on the training partition for each fold, therefore optimizing the performance of SVR on the training set.

4.5.3 Error Metrics

To measure the quality of the blending models, the metrics Relative Root Mean Square Error (rRMSE) and Relative Mean Absolute Error (rMAE) have been used. These are equations 3.18 and 3.19 as presented in Chapter 3 Section 3.2.3 represent the rRMSE and rMAE respectively.

4.6 Experimental results

In this section the methods proposed in Sections 4.1 and 4.1.3 are evaluated empirically for the four stations: Jaen, Lisbon, Madrid, and Seville. Physical and machine learning forecasts are compared and analysed. An experimental framework is proposed for the analysis of the integration problem in the following sections. Given that there is no standard validation methodology in place for the integration problem, common procedures for machine learning validation are proposed with several variations.

4.6.1 Results for local blending approaches

The empirical results are shown for the Horizon and General blending approaches described in section 4.1, and for the four stations: Jaen, Lisbon, Madrid, and Seville. Both approaches have been estimated using SVR with Linear and Radial kernels, resulting in four blending models, referred to as: LSVR-Horizon, RSVR-Horizon, LSVR-General, and RSVR-General. They are compared to the performance of the four original predictors. The average errors (rRMSE and rMAE) over all horizons and for each station (Jaen, Lisbon, Madrid, Seville) are shown in Tables 4.3 and 4.4 for GHI and DNI, respectively.

As it is observed in Table 4.3 the average error for GHI of the four predictors is consistently higher than any of the blending approaches. The improvement attained using the blending models, compared to the best performing predictor model, is maximum at Lisbon. At this station the rRMSE of the best predictor is 49.97% and LSVR-Horizon model value is 41.38%, i.e., an improvement of 8.59% (relative 17.19%). Jaen and Seville stations show a similar improvement of 6.04% (relative 17.32%) and 5.14% (relative 15.88%). Finally, the improvement of Madrid compared with the best predictor is of 4.1% (relative 11.29%).

Comparing blending techniques, it is observed that the rRMSE for different approaches is very similar, although the LSVR-Horizon performs best in Jaen (28.83%), Lisbon (41.38%) and Madrid (32.20%). In Seville the lowest rRMSE of 27.23% is reached using the General approach and Linear SVM, although the differences in terms of rRMSE are very small.

When rMAE is evaluated, blending models also outperform the four predictors. In this case all stations show around 3% improvement. The maximum improvement is achieved at Jaen 2.75% (relative 14.1%) and Seville 2.83% (relative 15.09%). The improvements on Lisbon are Madrid are better on absolute improvement, but worse on relative improvement with 3.86% (relative 11.73%) and 3.11% (relative 13.39%) respectively.

The best blending approach is RSVR-General for all stations: Jaen (16.75%), Lisbon (29.02%), Madrid (20.12%) and Seville (15.92%). It is also observed that there is some variability across stations, errors in Lisbon tend to be higher than others while the error in Seville tends to be the lower than the rest.

A similar performance is observed for DNI forecasting (see Table 4.4). All error metrics are lower for the blending methods than the four predictors, and comparing the machine learning methods also show that they perform similarly. The main difference between GHI and DNI is that the errors are

Table 4.3: General summary of error metrics on GHI measurements by station. Results are shown in %

	rRMSE				rMAE			
GHI	Jaen	Lisbon	Madrid	Seville	Jaen	Lisbon	Madrid	Seville
Satellite	35.85	49.97	39.53	33.64	22.63	36.86	27.18	20.49
WRFsolar	35.35	51.11	36.30	32.37	21.93	33.95	23.78	18.75
SmartPersistence	34.87	50.10	38.70	34.23	19.50	32.88	23.23	18.86
CIADCast	35.36	50.76	41.61	33.82	22.78	36.69	29.68	21.23
RSVR-General	29.19	41.94	32.89	27.67	**16.75**	**29.02**	**20.12**	**15.92**
LSVR-General	28.98	41.67	32.33	**27.23**	17.40	29.43	20.85	16.55
RSVR-Horizon	28.97	42.31	32.50	27.86	17.64	30.35	20.79	17.08
LSVR-Horizon	**28.83**	**41.38**	**32.20**	27.28	17.34	29.51	20.80	16.43

Table 4.4: General summary of error metrics on DNI measurements by station. Results are shown in %

	rRMSE				rMAE			
DNI	Jaen	Lisbon	Madrid	Seville	Jaen	Lisbon	Madrid	Seville
Satellite	53.36	88.99	66.69	51.35	36.25	67.96	46.53	34.74
WRFsolar	54.36	89.21	67.86	48.72	36.64	63.33	46.18	32.48
SmartPersistence	52.49	87.41	62.60	51.56	29.74	57.14	37.35	29.14
CIADCast	55.21	90.95	70.73	54.46	38.24	69.17	50.99	37.37
RSVR-General	45.15	73.99	56.19	42.54	**27.09**	**50.51**	**35.50**	**25.14**
LSVR-General	46.79	78.35	57.04	42.80	28.90	55.05	36.54	27.41
RSVR-Horizon	**44.42**	**73.45**	55.75	42.08	27.52	51.43	36.27	25.97
LSVR-Horizon	44.78	74.50	**55.48**	**41.81**	28.55	54.76	36.33	27.15

overall much worse than for GHI (for example, rRMSE from Lisbon goes from 41.38% to 73.45% and rMAE goes from 29.02% to 50.51%).

For rRMSE, the best approach appears to be the Horizon one. For Jaen (44.24%) and Lisbon (73.45%) the RSVR performs best, and for Madrid (55.48%) and Seville (41.81%) the LSVR has the lowest error. The improvement attained using the blending models, compared to the best performing input models, is maximum at Lisbon, with an improvement of 13.96% (relative 15.97%). The other stations show lower improvements like Jaen with 8.07% (relative 15.37%), Madrid with 7.12% (relative 11.37%) and Seville with 9.75 % (relative 18.91%).

With respect to rMAE and DNI, the average results show that the RSVR-General approach outperforms all other models, as it was also observed for GHI (see Table 4.3)). The most important improvement versus the performance of predictors, is observed at the Lisbon station with an improvement of 6.63% (relative 11.6%). The other stations present lower improvements of 2.65% (relative 8.91%) at Jaen, 1.84% (relative 4.95%) for Madrid and 4% (relative 13.72%) at Seville.

Figure 4.4 displays the rRMSE for GHI along the horizons for the four stations. Each of these figures contains a series called "Optimal Line" which represents the minimum errors from the predictors at each prediction horizon, which would be the optimal selection of models. It is observed that the blending approaches outperform the optimal line for all forecasting horizons. Blending machine learning models reach relative improvements with respect

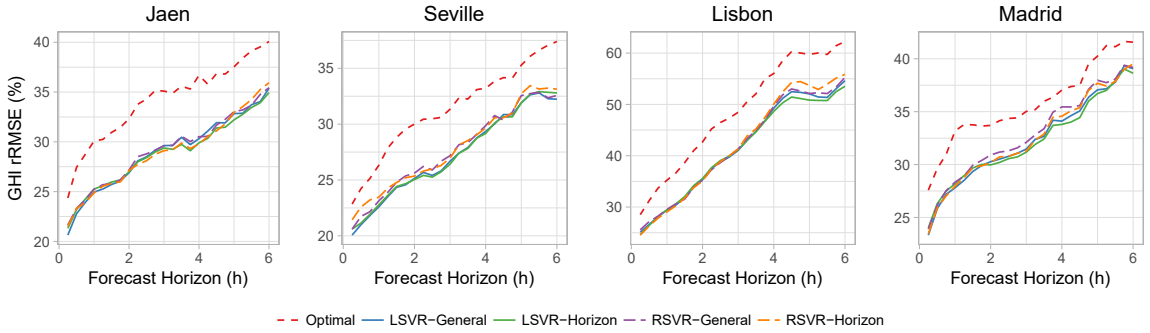

Figure 4.4: GHI rRMSE of blending models (General and Horizon) by horizon. "Optimal" displays the best performance out of the four predictors. Vertical values are in %. Note the difference range of values for the Lisbon station.

to the optimal line of 6% in the worst case and 18% in the best one. On the other hand, the errors of the blending machine learning models overlap with each other, being difficult to decide the best one approach for every horizon. However, LSVR-Horizon seems to perform better at least in some horizons. This is true for all stations, except for Seville, where this approach performs similar to LSVR-General model. In any case, the difference in terms of rRMSE is very small. The rRMSE in every station increases with the horizon, without large differences between blending models.

In Figure 4.5 the evolution of the rMAE for GHI is shown. In all stations, the optimal line is also worse than any other model, except at some horizons in Lisbon (from 330 to 360) and in Seville (from 285 to 345), where the optimal line is similar to the RSVR-Horizon approach. For rMAE, the relative improvement of blending models with respect to the optimal line goes from 3% (worst case) to 18% (best case, Jaen around 4 hours lead time). With respect to blending approaches, the best model for all stations is the RSVR-General approach. There is some overlap for some horizons with the LSVR-Horizon approach and with the LSVR-General approach in Lisbon station. The relative improvement of the RSVR-General model with respect to the LSVR-Horizon approach is around 3%, 4% and 5% in many cases. It is also observed that the RSVR-Horizon approach is very poor for rMAE metric. It ends up being outperformed by all other machine learning models at far horizons in Jaen (from h=285 to h=360) and Lisbon (from h=225 to h=360) and at most horizons for Seville station. For short horizons, all models start at the same range and the rMAE grows as the horizon increases, but the growth amount depends on the station.

Unlike the rRMSE, observing rMAE for GHI there is a clear best model for all stations, the RSVR-General approach. It is better for Jaen, Madrid and

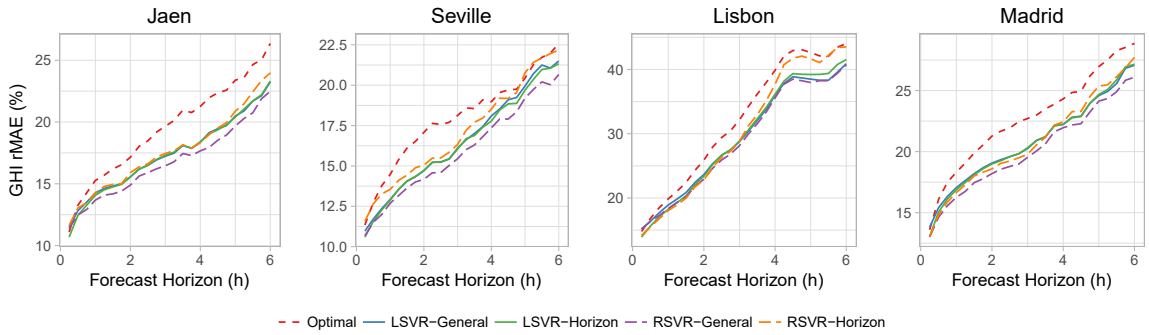

Figure 4.5: GHI rMAE of blending models (General and Horizon) by horizon. Vertical values are in %.

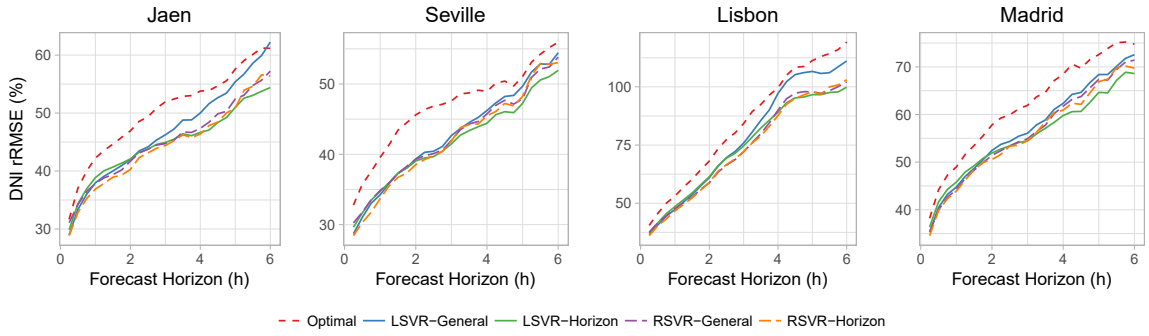

Figure 4.6: DNI rRMSE of blending models (General and Horizon) by horizon. Vertical values are in %.

Seville than the rest of the machine learning approaches, and it is similar to the LSVR-General and LSVR-Horizon approaches for Lisbon station.

Figure 4.6 shows the rRMSE for DNI forecasting along the horizons. Similarly to GHI forecasting, the blending approaches outperform the optimal line, with relative improvements between 5% and 16%. There is also overlap between blending machine learning models, but as the horizon increases, the LSVR-Horizon approach outperforms other models, where relative improvements of 2%, 3% and 4% are reached with respect to other approaches. At short horizons, the RSVR-Horizon model is slightly better, but the improvements are smaller. This pattern repeats for all stations, but it is more noticeable in Jaen, Madrid and Seville. Here the LSVR-General model is worse than the other models by a fair margin mainly for long horizons.

The Figure 4.7 shows the rMAE evolution for DNI forecasts. Unlike every other results presented before, here the optimal line outperforms or is similar to the blending models at short horizons (up to 60 minutes). As the horizon increases, the differences become greater and the RSVR-General approach appears to offer the best performance for long-term predictions, between 3% and 16% relative improvement over the optimal line. It is also observed that

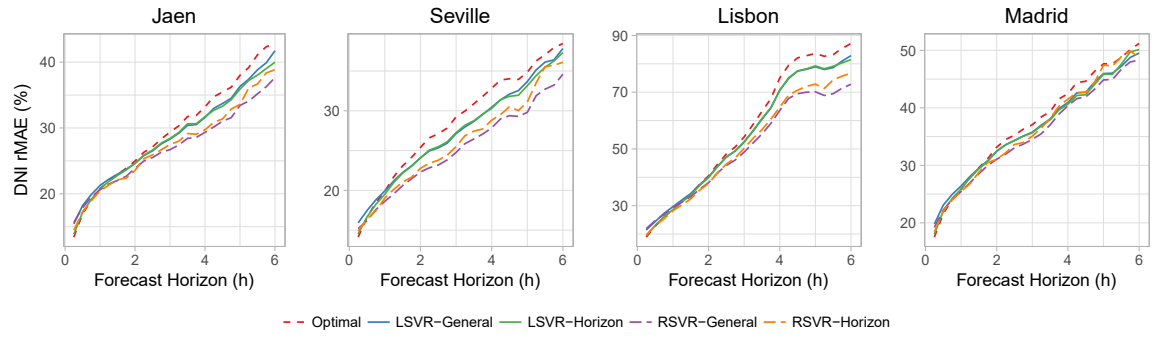

Figure 4.7: DNI rMAE of blending models (General and Horizon) by horizon. Vertical values are in %.

this approach overlaps with the RSVR-Horizon model in some stations and for short horizons. However, as the horizon increases, the advantage of the RSVR-General blending model becomes clearer, obtaining relative improvements up to 8%.

In summary, for GHI the four blending models performs similarly with respect to rRMSE, although for some horizons LSVR-Horizon is slightly better. For DNI and rRMSE, RSVR-Horizon model seems the best model for short horizons and the LSVR-Horizon one for far horizons. For rMAE the RSVR-General model is the best for both GHI and DNI, although for DNI this shows more clearly at far horizons.

It can also be seen that the differences between the "optimal" DNI model and blending models increases with the forecasting horizon, for rMAE. Therefore, it can be concluded that the main added value of the blended model is attained at the final of the forecasting horizon and that the improvement statistics summarized in Table 4.4, can mainly be attributed to the longest forecasting horizons. For the case of the GHI, the differences among forecasting horizons are not so relevant, and the blending models provide a similar improvement along the whole forecasting period.

4.6.2 Results for regional blending approaches

Both regional approaches are tested using the RSVR-General technique, the best local approach. Experimental results for the two regional approaches described in Section 4.1.3 are presented. In this case, SVR with Radial kernel has been used and both regional approaches are referred to as RSVR-General Mean (the first one: mean of local RSVR-General models) and RSVR-General Regional (the second one: a model constructed with all available predictors of all local stations). Other approaches (General linear or Horizon) might have been used, but only the General one has been selected here for empirical

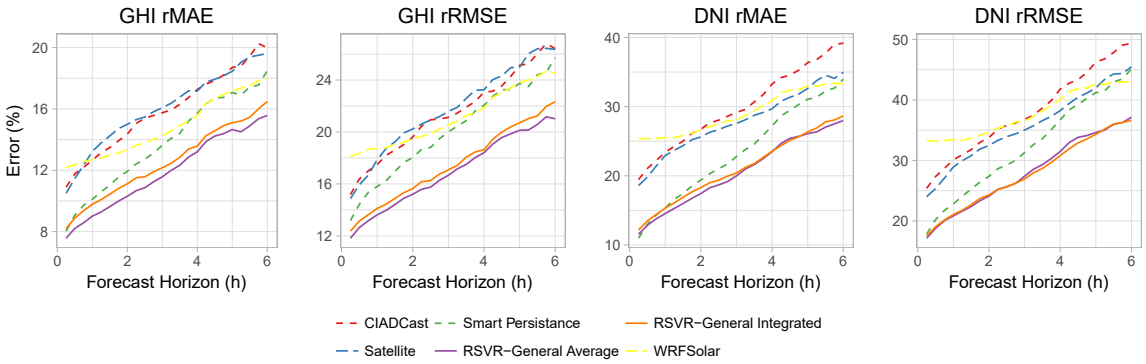

Figure 4.8: Metrics of aggregation models (Regional and Mean) by horizon. Vertical values are in %.

Table 4.5: General summary of error metrics on aggregated models for GHI and DNI. Shown in %

	GHI		DNI	
	rRMSE	rMAE	rRMSE	rMAE
Satellite	21.76	16.06	35.64	27.87
WRFsolar	21.10	14.83	37.78	29.09
SmartPersistence	20.03	13.78	32.29	23.44
CIADCast	21.43	15.89	38.03	30.00
RSVR-General Mean	**16.94**	**11.87**	28.01	**20.57**
RSVR-General Integrated	17.48	12.51	**27.84**	21.05

validation because it has shown good performance overall.

The results for both approaches are shown in Figure 4.8, where the metrics (rRMSE and rMAE) have been calculated at each horizon and for all irradiances (GHI and DNI). With the purpose of comparison, the figure also includes for each model, its mean over the four locations. For instance, CIADCast in Figure 8 is the mean of the CIADCast predictor over Jaen, Seville, Madrid, and Lisbon. As it can be observed, the two regional blending approaches outperform the mean of predictors for all horizons for both GHI and DNI. Also, the two regional approaches display similar performances. Table 4.5 shows the average rRMSE and rMAE for all horizons. It can be seen that, indeed, the two regional approaches are very close, but RSVR-General Mean display smaller errors (except for rRMSE DNI).

Improvements are observed to be on par with the local stations, with an improvement for GHI rRMSE of 3.09% (relative 15.43%) and rMAE of 2.96% (relative 19.96%). On DNI the improvement for rRMSE is 4.45% (relative 13.78%) and 2.87% (relative 12.24%) for rMAE.

As may be expected, averaged forecasts show considerably smaller rRMSE and rMAE than forecasts of single stations (Figures 4.4 to 4.7), because of the spatial decorrelation of the forecasts errors. An interesting feature of Figure 4.8 is that the reduction in the errors obtained by blending is higher than for the individual stations.

The comparison of Figure 4.8 and Figures 4.4 to 4.7 reveals some interesting features. When compared to local station results, the results for regional aggregation follow similar patterns on overall performance and horizon performance, where both machine learning methods far outperform the other predictors. In general, only when using blended models, the regional forecasts errors are considerable lower that the error for individual stations, which is clearly shown in the mentioned figures.

As with local stations, smart persistence manages to stay competitive in the shortest horizons, outperforming blending methods. For aggregation purposes Smart Persistence is the predictor model that takes the most advantage of the averaging of individual values, making it the best performing method out of the four predictor models. This happens on most horizons, but it is clearer on horizons from 15 minutes to 3 hours.

Comparing values in Tables 4.3, 4.4 and 4.5 the rRMSE values for the GHI reduces (mean over the forecasting period) from 27.67% (for Seville Table 4.3) to 16.94%, i.e., an improvement of 10.75% (relative 38.78%). Similarly, the rRMSE values for the DNI reduces from 41.81% (for Seville Table 4.4) to 27.84%, an improvement of 13.97% (relative 33.41%). Therefore, error reduction for DNI is considerable higher than for the GHI.

If rMAE is evaluated in the same way, the error for GHI is reduced from 15.92% (for Seville Table 4.3) to 11.87% an improvement of 4.05% (relative 25.44%) and for DNI it is reduced from 25.14% (for Seville Table 4.4) to 20.57% for an improvement of 3.57% (relative 14.79%). Therefore, it seems that, for regional forecast, the model blending provide superior performances that for individual stations. In other words, it seems that regional model blending takes advantage from the spatial decorrelation of the forecasting errors between models.

4.7 Summary of experimental conclusions

Various blending models are proposed throughout this chapter. Results in this section show the improvement of forecast blending over individual predictors. This is observed at all levels, for both, General and Horizon approaches; linear and radial SVR; GHI and DNI; and local and regional integration models. Furthermore, the improvement is observed for every prediction horizon independently of the method.

Local blending approaches all have similar performances, specially when rRMSE is measured. On the other hand, although similar, there are some approaches that perform better when rMAE is measured. For rRMSE the

best model on most horizons is the Horizon approach with the linear SVR algorithm, and for rMAE, it is the General approach with radial SVR. This happens on GHI and DNI measures alike.

Regional blending approaches also behave similarly. If GHI is measured, on both metrics, the average approach performs best. For DNI there is no clear best model as the average approach performs best on rRMSE and the integrated approach, on rMAE. There is a lot of overlap on the performance of the predictors on DNI.

The integration models are better at predicting GHI, achieving low error metrics than on DNI measurements. Also, the improvement that integration provides is higher for GHI.

To conclude this chapter, the main conclusion that can be drawn from these results is the beneficial impact of integration in forecasting. As the experimentation shows, it is always desirable to integrate irradiance forecasts when there are available predictors.

Chapter 5

Conclusions and future work

This chapter includes the final remarks of this thesis, including the conclusions drawn from the results, the possible future lines of research and a summary of the contributions.

5.1 Conclusions

Many challenges are ahead for renewables to achieve high participation shares in the energy market. This thesis has been conducted to frame and solve some challenges within the field of solar energy, and as empirical results show, the research objectives set have been satisfactorily met.

The first issue analysed has been the cloud classification problem. A random forest model is used to improve upon existing techniques by combining novel ceilometer information with common image metrics such as textural features, spectral features or cloud coverage. The main focus of the study is the ceilometer inputs, namely cloud height, thickness, and layers; information which has never seen use in the literature before. The influence of ceilometer has been exhaustively tested against difficult conditions by the standards in the literature, such as the 10-class cloud classification or the inclusion of multicloud examples.

As explored in Section 3.1, ceilometer information greatly improves the classification capabilities of camera information, with absolute improvements of 20.5% (relative 28.86%) in accuracy for the 10-class with multicloud sky type, which by current standards, is a very difficult problem. Also, it shows improvements in almost all cloud types, displaying greater improvements on partially and totally overcast skies. Overall, the results show the clear advantages of using ceilometer information for cloud classification. In all cases examined the use of ceilometer information proves to have a deep positive effect on all evaluated metrics.

The next work studies the irradiance estimation problem directly from camera images. A convolutional network model is used, combining multiple camera perspectives from the sky. Additionaly, this model has proven to benefit from extending the number of channels with additional information. Domain specific colour channels have been added with information about cloudy pixels and the pixel distance to the sun centre. These models aim to improve the overall capabilities of irradiance estimation by means of CNNs, multiple perspectives and domain-specific image information. Comparisons have been drawn against two baseline models, a cloud fraction model for statistical estimation (CF) and a random forest model for machine learning estimation (RF). More comparisons are made against various CNN designs.

Convolutional networks prove to be very powerful for irradiance estimation, clearly outperforming the baseline in every scenario. Both baselines offer worse estimation error than the CNN, and the alternatives evaluated prove the reasoning behind the final architecture designed. Furthermore, multiple camera perspectives outperform single camera results by a 1.21% (relative 8.48%) on rMAE and 0.44% (relative 2.38%) for rRMSE. All alternatives shown underperform compared to the final architecture. Additionally, if the results are broken down by sky type, the final three camera architecture is the best performing design on the majority of cloud types. The use of multiple camera perspectives with improved images is concluded to be highly beneficial to irradiance estimation on all evaluated metrics. CNNs in general perform well in this given task.

Lastly the forecasting problem is examined by means of forecast model integration. There are two different proposals for blending, a station-level and region-level integration model both using SVR. The station blending model combines information from a diverse set of four forecasting models to improve upon these predictors. This study has been made on GHI and DNI. Two different approaches have been tested, a horizon approach building a model for each available forecasting horizon, and a general approach that builds a single model for all horizons. Both these approaches have been tested with a linear and a non-linear SVR. The performance of the machine learning blending models has been exhaustively tested against the predictor models on each horizon.

The regional blending model used combines information from four stations from the southern region of the Iberian peninsula, which present distinct climatological behaviours. Two regional blending approaches are proposed and tested. The first is the average approach, that combines the integrated models from the stations with an average operation. The second is the inte-

grated approach, training another SVR model with all the forecasting models from the four stations simultaneously. The performance is tested against the predictors used for GHI and DNI forecasts.

Results have shown a clear improvement through the use of integration. All blending approaches are better than any of the individual forecasting models, on every horizon. Behaviours among machine learning methods are similar with subtle differences. These conclusions are seen both for GHI and DNI measurements. On GHI absolute improvements against the optimal combination of initial forecasting models are as high as 6.04% (relative 17.79%) for rRMSE, while for rMAE is 2.75% (relative 14.1%). On DNI absolute improvements are also high, with 13.96% (relative 15.97%) on rRMSE and 6.63% (relative 11.6%).

Observations made on local stations are similar to what happens with the regional models. Again, the integrated models outperform every predictor used by a considerable margin on both DNI and GHI. Both methods of regional integration perform in similar ways, improving upon the individual predictor models at every horizon. After studying both station and regional blending models it is firmly concluded that integration for short term forecasting is a powerful approach and that the used models are greatly beneficial to the forecasting of solar irradiance.

5.2 Future work

The research conducted in this thesis opens new venues for further investigation. In the cloud classification problem, an interesting alternative to the model proposed is the usage of CNNs. It is not trivial how the ceilometer information could be combined with the CNN. As the results of the irradiance estimation with CNNs section show, the usage of CNNs is apt for sky images, and as such a promising cloud classification model could be achieved.

As it has been observed with the cloud classification and irradiance estimation problems, a completely balanced dataset of cloud images is rare. For example the cloud classification capabilities of the proposed model could have been hindered by this misrepresentation of clouds. The application of data upsampling (or downsampling) could prove to generate a more robust dataset to learn with, which should be further researched. A hypothetical enhanced classification dataset could be built by applying SMOTE to the used dataset.

An interesting alternative for cloud classification appears, as ceilometer information is difficult to gather due to the high costs associated. To replace the ceilometer, sky images from different points of view could also be used to

gather data about the thickness and height of a cloud example. Using similar methods (Random Forests or CNNs) this information could be automatically detected by machine learning algorithms.

Regarding irradiance estimation, new lines of research appear. Given that domain specific colour channels improve the estimation CNN capabilities, further additional data could be included in the CNN either bi-dimensional (cloud motion vectors, heat maps, etc) or single values (wind speed, solar hour, date, ...). Including single values into CNNs is not trivial and has to be carefully considered, but with proper design it could hypothetically raise the accuracy of irradiance estimations.

This work shows the strength of using integration independently of historical values for short-term horizons. Using historical data from current and past measurements together with past forecasts could help the model detect further synergies. The experimentation conducted has focused at exploring the capabilities of blending only predictor models, not the influence of other meteorological information, so this line of research remains open for exploration. Most works in the literature combine a forecasting model with some meteorological data to improve forecasts, similar or higher improvements could be found if the presented blending models were combined with other meteorological data.

Furthermore, as of now, the models only receive information about forecasts for the specific forecasting horizon desired. By including information about forecasting models at other horizons simultaneously the error may decrease. It is hypothesized that forecasts from before and after the forecasting horizon desired are relevant to the desired forecasting horizon and may enhance the blending models.

As of right now, these blending techniques have been tested for short-term horizons up to 6 hours. This maximum horizon of 6 hours could be increased. By including data from forecasting models at further horizons, the integration technique presented may be able to combine more models. The exploration of horizons further than 6 hours remains untested, and is an interesting avenue for further research. This may not be a direct application of the approaches presented in this thesis and more sophisticated blending models may be needed to integrate data further than 6 hours.

5.3 Contributions

The contributions found across this thesis are machine learning applications that, through novel approaches adapted to the problem, significantly improve

traditional physical techniques. These accomplishments successfully achieve to satisfy the described objectives in the introductory Chapter 1.

First, a random forest model for cloud classification is proposed, where a novel approach that combines image and ceilometer information is combined. This instrument has never been used in the literature for automatic cloud classification and the beneficial influence of cloud thickness and height information is proven with this machine learning model.

Second, a CNN model is proposed for irradiance estimation with sky images. The network makes use of different camera perspectives and domain specific colour channels. This novel approach has never been used in the literature and improves estimation compared to two baseline methods, a statistical model (or CF) and feature-only (or RF) model.

Third and last, two different models for integration are contributed, one for local stations and another for regional forecasting, both using support vector regression models. The blending models combine forecasts from diverse sources to generate a final, more powerful prediction, which has been proven to significantly outperform all individual forecast models. Furthermore, the regional integration model combines forecasts from different solar stations to find a regional aggregate prediction, which, again, has been proven to outperform the average of individual predictors.

These contributions are significant in the field of solar irradiance, bringing relevant insights to irradiance estimation and forecasting, filling lacking parts from the literature.

Bibliography

[Adib et al., 2015] Adib, R., Murdock, H., Appavou, F., Brown, A., Epp, B., Leidreiter, A., Lins, C., Murdock, H., Musolino, E., Petrichenko, K., et al. (2015). Renewables 2015 global status report. *Paris: REN21 Secretariat*.

[Aguiar et al., 2016] Aguiar, L. M., Pereira, B., Lauret, P., Díaz, F., and David, M. (2016). Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. *Renewable Energy*, 97:599–610.

[Alonso-Montesinos and Batlles, 2015] Alonso-Montesinos, J. and Batlles, F. J. (2015). The use of a sky camera for solar radiation estimation based on digital image processing. *Energy*, 90(P1):377–386.

[Alonso-Montesinos et al., 2015a] Alonso-Montesinos, J., Batlles, F. J., and Bosch, J. L. (2015a). Beam, diffuse and global solar irradiance estimation with satellite imagery. *Energy Conversion and Management*, 105:1205–1212.

[Alonso-Montesinos et al., 2015b] Alonso-Montesinos, J., Batlles, F. J., and Portillo, C. (2015b). Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. *Energy Conversion and Management*, 105:1166–1177.

[Arbizu-Barrena et al., 2015] Arbizu-Barrena, C., Pozo-Vázquez, D., Ruiz-Arias, J. A., and Tovar-Pescador, J. (2015). Macroscopic cloud properties in the WRF NWP model: An assessment using sky camera and ceilometer data. *Journal of Geophysical Research*, 120(19):10297–10312.

[Arbizu-Barrena et al., 2017] Arbizu-Barrena, C., Ruiz-Arias, J. A., Rodríguez-Benítez, F. J., Pozo-Vázquez, D., and Tovar-Pescador, J. (2017). Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. *Solar Energy*, 155:1092–1103.

[Auligné, 2014a] Auligné, T. (2014a). Multivariate minimum residual method for cloud retrieval. part i: Theoretical aspects and simulated observation experiments. *Monthly Weather Review*, 142(12):4383–4398.

[Auligné, 2014b] Auligné, T. (2014b). Multivariate minimum residual method for cloud retrieval. part ii: Real observations experiments. *Monthly Weather Review*, 142(12):4399–4415.

[Bakirci, 2015] Bakirci, K. (2015). Models for the estimation of diffuse solar radiation for typical cities in turkey. *Energy*, 82:827–838.

[Bernecker et al., 2014] Bernecker, D., Riess, C., Angelopoulou, E., and Hornegger, J. (2014). Continuous short-term irradiance forecasts using sky images. *Solar Energy*, 110:303–315.

[Beyer et al., 1994] Beyer, H., Costanzo, C., Heinemann, D., and Reise, C. (1994). Short range forecast of pv energy production using satellite image analysis. In *Proc. 12th European Photovoltaic Solar Energy Conference*, volume 11, page 15. Amsterdam.

[Boers et al., 2010] Boers, R., De Haij, M. J., Wauben, W. M. F., Baltink, H. K., Van Ulft, L. H., Savenije, M., and Long, C. N. (2010). Optimized fractional cloudiness determination from five ground-based remote sensing techniques. *Journal of Geophysical Research Atmospheres*, 115(24).

[Bosch et al., 2008] Bosch, J., López, G., and Batlles, F. (2008). Daily solar irradiation estimation over a mountainous area using artificial neural networks. *Renewable Energy*, 33(7):1622–1628.

[Breiman, 2001] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

[Brouwer et al., 2014] Brouwer, A. S., Van Den Broek, M., Seebregts, A., and Faaij, A. (2014). Impacts of large-scale intermittent renewable energy sources on electricity systems, and how these can be modeled. *Renewable and Sustainable Energy Reviews*, 33:443–466.

[Calbó and Sabburg, 2008] Calbó, J. and Sabburg, J. (2008). Feature extraction from Whole-sky ground-based images for cloud-type recognition. *Journal of Atmospheric and Oceanic Technology*, 25(1):3–14.

[Carslaw et al., 2002] Carslaw, K., Harrison, R., and Kirkby, J. (2002). Cosmic rays, clouds, and climate. *Science*, 298(5599):1732–1737.

[Caruana et al., 2008] Caruana, R., Karampatziakis, N., and Yessenalina, A. (2008). An empirical evaluation of supervised learning in high dimensions. In *Proceedings of the 25th International Conference on Machine Learning*, pages 96–103.

[Caruana and Niculescu-Mizil, 2006] Caruana, R. and Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In *ACM International Conference Proceeding Series*, volume 148, pages 161–168.

[Castro-Díez et al., 2002] Castro-Díez, Y., Pozo-Vázquez, D., Rodrigo, F., and Esteban-Parra, M. (2002). Nao and winter temperature variability in southern europe. *Geophysical Research Letters*, 29(8):1–1.

[Cazorla et al., 2008] Cazorla, A., Olmo, F. J., and Alados-Arbotedas, L. (2008). Development of a sky imager for cloud cover assessment. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 25(1):29–39.

[Chen et al., 2012] Chen, L., Yan, G., Wang, T., Ren, H., Calbó, J., Zhao, J., and McKenzie, R. (2012). Estimation of surface shortwave radiation components under all sky conditions: Modeling and sensitivity analysis. *Remote Sensing of Environment*, 123:457–469.

[Cheng and Lin, 2017] Cheng, H. . and Lin, C. . (2017). Cloud detection in all-sky images via multi-scale neighborhood features and multiple supervised learning techniques. *Atmospheric Measurement Techniques*, 10(1):199–208.

[Cheng and Yu, 2015] Cheng, H. . and Yu, C. . (2015). Block-based cloud classification with statistical features and distribution of local texture features. *Atmospheric Measurement Techniques*, 8(3):1173–1182.

[Chow et al., 2011] Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., and Washom, B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, 85(11):2881–2893.

[Chu et al., 2013] Chu, Y., Pedro, H. T., and Coimbra, C. F. (2013). Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning. *Solar Energy*, 98:592–603.

[Cloud Appreciation Society, 2011] Cloud Appreciation Society (2011).

[Coppolino, 1994] Coppolino, S. (1994). A new correlation between clearness index and relative sunshine. *Renewable Energy*, 4(4):417–423.

[Costa-Surós et al., 2014] Costa-Surós, M., Calbó, J., González, J. A., and Long, C. N. (2014). Comparing the cloud vertical structure derived from several methods based on radiosonde profiles and ground-based remote sensing measurements. *Atmospheric Measurement Techniques*, 7(8):2757–2773.

[Crammer and Singer, 2001] Crammer, K. and Singer, Y. (2001). On the algorithmic implementation of multiclass kernel-based vector machines. *Journal of Machine Learning Research*, 2(2):265–292.

[Davies et al., 1975] Davies, J. A., Schertzer, W., and Nunez, M. (1975). Estimating global solar radiation. *Boundary-Layer Meteorology*, 9(1):33–52.

[de Coninck et al., 2018] de Coninck, H., Revi, A., Babiker, M., Bertoldi, P., Buckeridge, M., Cartwright, A., Dong, W., Ford, J., Fuss, S., Hourcade, J.-C., and et al. (2018). Strengthening and implementing the global response. In *Global Warming of 1.5C an IPCC special report on the impacts of global warming of 1.5C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change*, chapter 4, page 132. Intergovernmental Panel on Climate Change.

[De Gooijer and Hyndman, 2006] De Gooijer, J. G. and Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22(3):443–473.

[Deng et al., 2014] Deng, A., Gaudet, B., Dudhia, J., and Alapaty, K. (2014). 12.5 implementation and evaluation of a new shallow convection scheme in wrf. In *94th American Meteorological Society Annual Meeting, 26th Conference on Weather Analysis and Forecasting/ 22nd Conference on Numerical Weather Prediction*.

[Dersch et al., 2019] Dersch, J., Schroedter-Homscheidt, M., Gairaa, K., Hanrieder, N., Landelius, T., Lindskog, M., Müller, S., Ramirez Santigosa, L., Sirch, T., and Wilbert, S. (2019). Impact of dni nowcasting on annual revenues of csp plants for a time of delivery based feed in tariff. *Meteorologische Zeitschrift*.

[Descombes et al., 2014] Descombes, G., Auligné, T., Lin, H.-C., Xu, D., Schwartz, C., and Vandenberghe, F. (2014). Multi-sensor advection diffusion nowcast (madcast) for cloud analysis and short-term prediction. *Nat. Center Atmos. Res., Boulder, CO, USA, Tech. Note NCAR/TN-509+ STR*.

[Dev et al., 2016] Dev, S., Savoy, F. M., Lee, Y. H., and Winkler, S. (2016). Estimation of solar irradiance using ground-based whole sky imagers. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 7236–7239. IEEE.

[Drucker et al., 1997] Drucker, H., Burges, C. J., Kaufman, L., Smola, A. J., and Vapnik, V. (1997). Support vector regression machines. In *Advances in neural information processing systems*, pages 155–161.

[Ela et al., 2011] Ela, E., Milligan, M., and Kirby, B. (2011). Operating reserves and variable generation. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

[Şenkal and Kuleli, 2009] Şenkal, O. and Kuleli, T. (2009). Estimation of solar radiation over turkey using artificial neural network and satellite data. *Applied Energy*, 86(7):1222–1228.

[Fan et al., 2018a] Fan, J., Wang, X., Wu, L., Zhang, F., Bai, H., Lu, X., and Xiang, Y. (2018a). New combined models for estimating daily global solar radiation based on sunshine duration in humid regions: A case study in south china. *Energy Conversion and Management*, 156:618–625.

[Fan et al., 2018b] Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., Lu, X., and Xiang, Y. (2018b). Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in china. *Energy Conversion and Management*, 164:102–111.

[Fu and Cheng, 2013] Fu, C.-L. and Cheng, H.-Y. (2013). Predicting solar irradiance with all-sky image features via regression. *Solar Energy*, 97:537–550.

[Gauchet et al., 2012] Gauchet, C., Blanc, P., Espinar, B., Charbonnier, B., and Demengel, D. (2012). Surface solar irradiance estimation with low-cost fisheye camera. *Workshop on "Remote Sensing Measurements for Renewable Energy"*, page 5.

[Ghonima et al., 2012] Ghonima, M. S., Urquhart, B., Chow, C. W., Shields, J. E., Cazorla, A., and Kleissl, J. (2012). A method for cloud detection and opacity classification based on ground based sky imagery. *Atmospheric Measurement Techniques*, 5(11):2881–2892.

[Ghorbani et al., 2016] Ghorbani, M. A., Khatibi, R., FazeliFard, M. H., Naghipour, L., and Makarynskyy, O. (2016). Short-term wind speed predictions with machine learning techniques. *Meteorology and Atmospheric Physics*, 128(1):57–72.

[Guillot et al., 2012] Guillot, E. M., Vonder Haar, T. H., Forsythe, J. M., and Fletcher, S. J. (2012). Evaluating satellite-based cloud persistence and displacement nowcasting techniques over complex terrain. *Weather and Forecasting*, 27(2):502–514.

[Haiden et al., 2015] Haiden, T., Forbes, R., Ahlgrimm, M., and Bozzo, A. (2015). The skill of ECMWF cloudiness forecasts. *ECMWF Newsl.*, 143:14–19.

[Hamann, 2017] Hamann, H. F. (2017). A multi-scale, multi-model, machine-learning solar forecasting technology. Technical report, USDOE Office of Energy Efficiency and Renewable Energy (EERE).

[Haupt et al., 2017] Haupt, S. E., Kosović, B., Jensen, T., Lazo, J. K., Lee, J. A., Jiménez, P. A., Cowie, J., Wiener, G., McCandless, T. C., Rogers, M., Miller, S., Sengupta, M., Xie, Y., Hinkelman, L., Kalb, P., and Heiser, J. (2017). Building the sun4cast system: Improvements in solar power forecasting. *Bulletin of the American Meteorological Society*, 99(1):121–136.

[Heinle et al., 2010] Heinle, A., Macke, A., and Srivastav, A. (2010). Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques*, 3(3):557–567.

[Hoegh-Guldberg et al., 2018] Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., and et al. (2018). Impacts of 1.5c of global warming on natural and human systems. In *Global Warming of 1.5C an IPCC special report on the impacts of global warming of 1.5C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change*, chapter 3, page 138. Intergovernmental Panel on Climate Change.

[Houze Jr, 1993] Houze Jr, R. A. (1993). Cloud dynamics. *Cloud dynamics*.

[Hoyt, 1978] Hoyt, D. V. (1978). Interannual cloud-cover variations in the contiguous united states. *Journal of Applied Meteorology*, 17(3):354–357.

[Huang et al., 2013] Huang, J., Korolkiewicz, M., Agrawal, M., and Boland, J. (2013). Forecasting solar radiation on an hourly time scale using a coupled AutoRegressive and dynamical system (CARDS) model. *Solar Energy*, 87:136–149.

[Iacono et al., 2008] Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the aer radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13).

[Illingworth et al., 2007] Illingworth, A. J., Hogan, R. J., O'Connor, E. J., Bouniol, D., Brooks, M. E., Delanoë, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W. F., Haeffelin, M., Klein Baltinik, H., Krasnov, O. A., Pelon, J., Piriou, J. ., Protat, A., Russchenberg, H. W. J., Seifert, A., Tompkins, A. M., van Zadelhoff, G. ., Vinit, F., Willen, U., Wilson, D. R., and Wrench, C. L. (2007). Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bulletin of the American Meteorological Society*, 88(6):883–898.

[Jayadevan et al., 2015] Jayadevan, V. T., Rodriguez, J. J., and Cronin, A. D. (2015). A new contrast-enhancing feature for cloud detection in ground-based sky images. *Journal of Atmospheric and Oceanic Technology*, 32(2):209–219.

[Jiménez et al., 2012] Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., and García-Bustamante, E. (2012). A revised scheme for the wrf surface layer formulation. *Monthly Weather Review*, 140(3):898–918.

[Jimenez et al., 2015] Jimenez, P. A., Hacker, J. P., Dudhia, J., Haupt, S. E., Ruiz-Arias, J. A., Gueymard, C. A., Thompson, G., Eidhammer, T., and Deng, A. (2015). WRF-solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction. *Bulletin of the American Meteorological Society*, 97(7):1249–1264.

[Kabir et al., 2018] Kabir, E., Kumar, P., Kumar, S., Adelodun, A. A., and Kim, K.-H. (2018). Solar energy: Potential and future prospects. *Renewable and Sustainable Energy Reviews*, 82:894–900.

[Kashyap et al., 2015] Kashyap, Y., Bansal, A., and Sao, A. K. (2015). Solar radiation forecasting with multiple parameters neural networks. *Renewable and Sustainable Energy Reviews*, 49:825–835.

[Kaur et al., 2016] Kaur, A., Nonnenmacher, L., Pedro, H. T. C., and Coimbra, C. F. M. (2016). Benefits of solar forecasting for energy imbalance markets. *Renewable Energy*, 86:819–830.

[Kazantzidis et al., 2012] Kazantzidis, A., Tzoumanikas, P., Bais, A. F., Fotopoulos, S., and Economou, G. (2012). Cloud detection and classification with the use of whole-sky ground-based images. *Atmospheric Research*, 113:80–88.

[Kingma and Ba, 2014] Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. arXiv: 1412.6980.

[Kliangsuwan and Heednacram, 2015] Kliangsuwan, T. and Heednacram, A. (2015). Feature extraction techniques for ground-based cloud type classification. *Expert Systems with Applications*, 42(21):8294–8303.

[Koca et al., 2011] Koca, A., Oztop, H. F., Varol, Y., and Koca, G. O. (2011). Estimation of solar radiation using artificial neural networks with different input parameters for mediterranean region of anatolia in turkey. *Expert Systems with Applications*, 38(7):8756–8762.

[Kokhanovsky, 2006] Kokhanovsky, A. A. (2006). Cloud optics. *Atmospheric and Oceanographic Sciences Library*, 34:123.

[Kreuter and Blumthaler, 2013] Kreuter, A. and Blumthaler, M. (2013). Feasibility of polarized all-sky imaging for aerosol characterization. *Atmospheric Measurement Techniques*, 6(7):1845–1854.

[Kreuter et al., 2009] Kreuter, A., Zangerl, M., Schwarzmann, M., and Blumthaler, M. (2009). All-sky imaging: A simple, versatile system for atmospheric research. *Applied Optics*, 48(6):1091–1097.

[Kriebel et al., 2003] Kriebel, K. T., Gesell, G., Kästner, M., and Mannstein, H. (2003). The cloud analysis tool APOLLO: Improvements and validations. *International Journal of Remote Sensing*, 24(12):2389–2408.

[Kuhn, 2008] Kuhn, M. (2008). Caret package. *J.Stat.Softw.*, 28(5):1–26.

[Kühnert et al., 2013] Kühnert, J., Lorenz, E., and Heinemann, D. (2013). Chapter 11-satellite-based irradiance and power forecasting for the german energy market. *Solar Energy Forecasting and Resource Assessment*, pages 267–297.

[Lara-Fanego et al., 2012] Lara-Fanego, V., Ruiz-Arias, J., Pozo-Vázquez, D., Santos-Alamillos, F., and Tovar-Pescador, J. (2012). Evaluation of the wrf model solar irradiance forecasts in andalusia (southern spain). *Solar Energy*, 86(8):2200–2217.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436.

[Lee et al., 1990] Lee, J., Weger, R. C., Sengupta, S. K., and Welch, R. M. (1990). A Neural Network Approach to Cloud Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):846–855.

[Lei et al., 2009] Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4):915–920.

[Lew et al., 2011] Lew, D., Piwko, R., Jordan, G., Miller, N., Clark, K., Freeman, L., and Milligan, M. (2011). Western wind and solar integration study. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

[Li et al., 2011] Li, Q., Lu, W., and Yang, J. (2011). A hybrid thresholding algorithm for cloud detection on ground-based color images. *Journal of Atmospheric and Oceanic Technology*, 28(10):1286–1296.

[Li et al., 2016a] Li, Q., Zhang, Z., Lu, W., Yang, J., Ma, Y., and Yao, W. (2016a). From pixels to patches: A cloud classification method based on a bag of micro-structures. *Atmospheric Measurement Techniques*, 9(2):753–764.

[Li et al., 2014] Li, Y., Thompson, D. W. J., Stephens, G. L., and Bony, S. (2014). A global survey of the instantaneous linkages between cloud vertical structure and large-scale climate. *Journal of Geophysical Research*, 119(7):3770–3792.

[Li et al., 2016b] Li, Z., Rahman, S. M., Vega, R., and Dong, B. (2016b). A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting. *Energies*, 9(1):55.

[Liaw and Wiener, 2002] Liaw, A. and Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3):18–22.

[Liu et al., 2014] Liu, D., Niu, D., Wang, H., and Fan, L. (2014). Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renewable Energy*, 62:592–597.

[Long and DeLuisi, 1998] Long, C. N. and DeLuisi, J. J. (1998). Development of an automated hemispheric sky imager for cloud fraction retrievals. *10th Symposium on Meteorological Observations and Instrumentation*, pages 171–174.

[Long and Dutton, 2010] Long, C. N. and Dutton, E. G. (2010). Bsrn global network recommended qc tests, v2. x.

[Long et al., 2006] Long, C. N., Sabburg, J. M., Calbó, J., and Pagès, D. (2006). Retrieving cloud characteristics from ground-based daytime color all-sky images. *Journal of Atmospheric and Oceanic Technology*, 23(5):633–652.

[Lorenz and Heinemann, 2012] Lorenz, E. and Heinemann, D. (2012). Prediction of solar irradiance and photovoltaic power. *Comprehensive Renewable Energy*.

[Lorenz et al., 2014] Lorenz, E., Kühnert, J., and Heinemann, D. (2014). Overview of irradiance and photovoltaic power prediction. In *Weather matters for energy*, pages 429–454. Springer.

[Lorenz et al., 2016] Lorenz, E., Kühnert, J., Heinemann, D., Nielsen, K. P., Remund, J., and Müller, S. C. (2016). Comparison of global horizontal irradiance forecasts based on numerical weather prediction models with different spatio-temporal resolutions. *Progress in Photovoltaics: Research and Applications*, 24(12):1626–1640.

[Lu et al., 2015] Lu, S., Hwang, Y., Khabibrakhmanov, I., Marianno, F. J., Shao, X., Zhang, J., Hodge, B. M., and Hamann, H. F. (2015). Machine learning based multi-physical-model blending for enhancing renewable energy forecast - improvement via situation dependent error correction. In *2015 European Control Conference (ECC)*, pages 283–290.

[Mannstein et al., 2010] Mannstein, H., Brömser, A., and Bugliaro, L. (2010). Ground-based observations for the validation of contrails and cirrus detection in satellite imagery. *Atmospheric Measurement Techniques*, 3(3):655–669.

[Marquez and Coimbra, 2013] Marquez, R. and Coimbra, C. F. M. (2013). Intra-hour DNI forecasting based on cloud tracking image analysis. *Solar Energy*, 91:327–336.

[Martínez-Chico et al., 2011] Martínez-Chico, M., Batlles, F. J., and Bosch, J. L. (2011). Cloud classification in a mediterranean location using radiation data and sky images. *Energy*, 36(7):4055–4062.

[Martucci et al., 2010] Martucci, G., Milroy, C., and O'Dowd, C. D. (2010). Detection of cloud-base height using Jenoptik CHM15k and Vaisala CL31 ceilometers. *Journal of Atmospheric and Oceanic Technology*, 27(2):305–318.

[Mateos et al., 2014] Mateos, D., Antón, M., Valenzuela, A., Cazorla, A., Olmo, F. J., and Alados-Arboledas, L. (2014). Efficiency of clouds on shortwave radiation using experimental data. *Applied Energy*, 113:1216–1219.

[Mathiesen and Kleissl, 2011] Mathiesen, P. and Kleissl, J. (2011). Evaluation of numerical weather prediction for intra-day solar forecasting in the continental united states. *Solar Energy*, 85(5):967–977.

[Mellit, 2008] Mellit, A. (2008). Artificial intelligence technique for modelling and forecasting of solar radiation data: a review. *International Journal of Artificial Intelligence and Soft Computing*, 1(1):52.

[Miller et al., 2012] Miller, S., Rogers, M., Heidinger, A., Laszlo, I., and Sengupta, M. (2012). Cloud advection schemes for short-term satellite-based insolation forecasts. In *Proc. of World Renewable Energy Forum*.

[Miller et al., 2018] Miller, S. D., Rogers, M. A., Haynes, J. M., Sengupta, M., and Heidinger, A. K. (2018). Short-term solar irradiance forecasting via satellite/model coupling. *Solar Energy*, 168:102–117.

[Mohandes et al., 2004] Mohandes, M. A., Halawani, T. O., Rehman, S., and Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Renewable Energy*, 29(6):939–947.

[Mori and Chang, 2003] Mori, N. and Chang, K.-A. (2003). Introduction to mpiv. *user reference manual*, 14.

[Nakanishi and Niino, 2009] Nakanishi, M. and Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan. Ser. II*, 87(5):895–912.

[Nonnenmacher and Coimbra, 2014] Nonnenmacher, L. and Coimbra, C. F. (2014). Streamline-based method for intra-day solar forecasting through remote sensing. *Solar Energy*, 108:447–459.

[Orjuela-Cañón et al., 2017] Orjuela-Cañón, A. D., Hernández, J., and Rivero, C. R. (2017). Very short term forecasting in global solar irradiance using linear and nonlinear models. In *2017 IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA)*, pages 1–5.

[Ozgoren et al., 2012] Ozgoren, M., Bilgili, M., and Sahin, B. (2012). Estimation of global solar radiation using ANN over turkey. *Expert Systems with Applications*, 39(5):5043–5051.

[Perez et al., 2010] Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr, K., Renné, D., and Hoff, T. E. (2010). Validation of short and medium term operational solar radiation forecasts in the us. *Solar Energy*, 84(12):2161–2172.

[Pfister et al., 2003] Pfister, G., McKenzie, R. L., Liley, J. B., Thomas, A., Forgan, B. W., and Long, C. N. (2003). Cloud Coverage Based on All-Sky Imaging and Its Impact on Surface Solar Irradiance. *Journal of Applied Meteorology*, 42(10):1421–1434.

[Pierro et al., 2017] Pierro, M., De Felice, M., Maggioni, E., Moser, D., Perotto, A., Spada, F., and Cornaro, C. (2017). Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data. *Solar Energy*, 158:1026–1038.

[Pincus et al., 2011] Pincus, R., Hofmann, R. J. P., Anderson, J. L., Raeder, K., Collins, N., and Whitaker, J. S. (2011). Can fully accounting for clouds in data assimilation improve short-term forecasts by global models? *Monthly Weather Review*, 139(3):946–957.

[Pozo-Vazquez et al., 2011] Pozo-Vazquez, D., Santos-Alamillos, F. J., Lara-Fanego, V., Ruiz-Arias, J. A., and Tovar-Pescador, J. (2011). The impact of the nao on the solar and wind energy resources in the mediterranean area. In *Hydrological, Socioeconomic and Ecological Impacts of the North Atlantic Oscillation in the Mediterranean Region*, pages 213–231. Springer.

[Pozo-Vázquez et al., 2004] Pozo-Vázquez, D., Tovar-Pescador, J., Gámiz-Fortis, S., Esteban-Parra, M., and Castro-Díez, Y. (2004). Nao and solar radiation variability in the european north atlantic region. *Geophysical Research Letters*, 31(5).

[Quesada-Ruiz et al., 2014] Quesada-Ruiz, S., Chu, Y., Tovar-Pescador, J., Pedro, H. T. C., and Coimbra, C. F. M. (2014). Cloud-tracking methodology for intra-hour DNI forecasting. *Solar Energy*, 102:267–275.

[Rangarajan et al., 1984] Rangarajan, S., Swaminathan, M., and Mani, A. (1984). Computation of solar radiation from observations of cloud cover. *Solar Energy*, 32(4):553–556.

[Ren et al., 2015] Ren, Y., Suganthan, P. N., and Srikanth, N. (2015). Ensemble methods for wind and solar power forecasting—a state-of-the-art review. *Renewable and Sustainable Energy Reviews*, 50:82–91.

[Renné, 2014] Renné, D. S. (2014). Emerging meteorological requirements to support high penetrations of variable renewable energy sources: solar energy. In *Weather matters for energy*, pages 257–273. Springer.

[Rigollier et al., 2000] Rigollier, C., Bauer, O., and Wald, L. (2000). On the clear sky model of the esra—european solar radiation atlas—with respect to the heliosat method. *Solar energy*, 68(1):33–48.

[Rigollier et al., 2004] Rigollier, C., Lefèvre, M., and Wald, L. (2004). The method heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy*, 77(2):159–169.

[Rodríguez-Benítez et al., 2016] Rodríguez-Benítez, F. J., Arbizu-barrena, C., Quesada-Ruiz, S., Tovar-Pescador, J., and Pozo-Vazquez, D. (2016). A "smart" method for short-term solar radiation forecasting. In *16th EMS / 11th ECAC*.

[Rodríguez-Benítez et al., 2018] Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Santos-Alamillos, F. J., Tovar-Pescador, J., and Pozo-Vázquez, D. (2018). Analysis of the intra-day solar resource variability in the iberian peninsula. *Solar Energy*, 171:374–387.

[Ruiz-Arias et al., 2016] Ruiz-Arias, J. A., Arbizu-Barrena, C., Santos-Alamillos, F. J., Tovar-Pescador, J., and Pozo-Vázquez, D. (2016). Assessing the surface solar radiation budget in the wrf model: A spatiotemporal analysis of the bias and its causes. *Monthly Weather Review*, 144(2):703–711.

[Rumi et al., 2013] Rumi, E., Kerr, D., Coupland, J. M., Sandford, A. P., and Brettle, M. J. (2013). Automated cloud classification using a ground based infra-red camera and texture analysis techniques. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 8890.

[Sabbagh et al., 1977] Sabbagh, J. A., Sayigh, A. a. M., and El-Salam, E. M. A. (1977). Estimation of the total solar radiation from meteorological data. *Sol. Energy; (United States)*, 19:3.

[Sabburg and Wong, 1999] Sabburg, J. and Wong, J. (1999). Evaluation of a Ground-Based Sky Camera System for Use inSurface Irradiance Measurement. *Journal of Atmospheric and Oceanic Technology*, 16(6):752–759.

[Salazar et al., 2011] Salazar, E., Sansó, B., Finley, A. O., Hammerling, D., Steinsland, I., Wang, X., and Delamater, P. (2011). Comparing and blending regional climate model predictions for the american southwest. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(4):586–605.

[Salcedo-Sanz et al., 2014] Salcedo-Sanz, S., Casanova-Mateo, C., Pastor-Sánchez, A., and Sánchez-Girón, M. (2014). Daily global solar radiation prediction based on a hybrid coral reefs optimization – extreme learning machine approach. *Solar Energy*, 105:91–98.

[Santamaría-Bonfil et al., 2016] Santamaría-Bonfil, G., Reyes-Ballesteros, A., and Gershenson, C. (2016). Wind speed forecasting for wind farms: A method based on support vector regression. *Renewable Energy*, 85:790–809.

[Santos-Alamillos et al., 2012] Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., and Tovar-Pescador, J. (2012). Analysis of spatiotemporal balancing between wind and solar energy resources in the southern iberian peninsula. *Journal of Applied Meteorology and Climatology*, 51(11):2005–2024.

[Santos-Alamillos et al., 2014] Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., and Tovar-Pescador, J. (2014). A methodology for evaluating the spatial variability of wind energy resources: application to assess the potential contribution of wind energy to baseload power. *Renew Energy Sep*, 69:147e56.

[Sharma et al., 2016] Sharma, V., Yang, D., Walsh, W., and Reindl, T. (2016). Short term solar irradiance forecasting using a mixed wavelet neural network. *Renewable Energy*, 90:481–492.

[Siddiqui et al., 2019] Siddiqui, T. A., Bharadwaj, S., and Kalyanaraman, S. (2019). A deep learning approach to solar-irradiance forecasting in sky-videos. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 2166–2174. IEEE.

[Singh and Glennen, 2005] Singh, M. and Glennen, M. (2005). Automated ground-based cloud recognition. *Pattern Analysis and Applications*, 8(3):258–271.

[Smith, 1978] Smith, A. R. (1978). COLOR GAMUT TRANSFORM PAIRS. *Comput Graph (ACM)*, 12(3):12–19.

[Smola and Schölkopf, 2004] Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222.

[Stefferud et al., 2012] Stefferud, K., Kleissl, J., and Schoene, J. (2012). Solar forecasting and variability analyses using sky camera cloud detection amp; motion vectors. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–6.

[Su et al., 2015] Su, H., Maji, S., Kalogerakis, E., and Learned-Miller, E. (2015). Multi-view convolutional neural networks for 3d shape recognition. In *2015 IEEE International Conference on Computer Vision (ICCV)*, page 945–953. IEEE.

[Taravat et al., 2015] Taravat, A., Del Frate, F., Cornaro, C., and Vergari, S. (2015). Neural networks and support vector machine algorithms for auto-

matic cloud classification of whole-sky ground-based images. *IEEE Geoscience and Remote Sensing Letters*, 12(3):666–670.

[Tascikaraoglu and Uzunoglu, 2014] Tascikaraoglu, A. and Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renewable and Sustainable Energy Reviews*, 34:243–254.

[Thompson and Eidhammer, 2014] Thompson, G. and Eidhammer, T. (2014). A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the atmospheric sciences*, 71(10):3636–3658.

[Torres et al., 2019] Torres, J. F., Troncoso, A., Koprinska, I., Wang, Z., and Martínez□ Álvarez, F. (2019). Big data solar power forecasting based on deep learning and multiple data sources. *Expert Systems*, page e12394.

[Trigo et al., 2002] Trigo, R. M., Osborn, T. J., and Corte-Real, J. M. (2002). The north atlantic oscillation influence on europe: climate impacts and associated physical mechanisms. *Climate research*, 20(1):9–17.

[Trigo et al., 2004] Trigo, R. M., Pozo-Vázquez, D., Osborn, T. J., Castro-Díez, Y., Gámiz-Fortis, S., and Esteban-Parra, M. J. (2004). North atlantic oscillation influence on precipitation, river flow and water resources in the iberian peninsula. *International Journal of Climatology*, 24(8):925–944.

[Tuohy et al., 2015] Tuohy, A., Zack, J., Haupt, S. E., Sharp, J., Ahlstrom, M., Dise, S., Grimit, E., Mohrlen, C., Lange, M., Casado, M. G., Black, J., Marquis, M., and Collier, C. (2015). Solar Forecasting: Methods, Challenges, and Performance. *IEEE Power and Energy Magazine*, 13(6):50–59.

[Tzoumanikas and Kazantzidis, 2016] Tzoumanikas, P. and Kazantzidis, A. (2016). Cloud detection and properties with whole sky images. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, page 19.

[Tzoumanikas et al., 2016] Tzoumanikas, P., Nikitidou, E., Bais, A. F., and Kazantzidis, A. (2016). The effect of clouds on surface solar irradiance, based on data from an all-sky imaging system. *Renewable Energy*, 95:314–322.

[Vázquez-Cuervo et al., 2004] Vázquez-Cuervo, J., Armstrong, E. M., and Harris, A. (2004). The effect of aerosols and clouds on the retrieval of infrared sea surface temperatures. *Journal of Climate*, 17(20):3921–3933.

[Viúdez-Mora et al., 2015] Viúdez-Mora, A., Costa-Surós, M., Calbó, J., and González, J. A. (2015). Modeling atmospheric longwave radiation at the

surface during overcast skies: The role of cloud base height. *Journal of Geophysical Research*, 120(1):199–214.

[Vislocky and Fritsch, 1995] Vislocky, R. L. and Fritsch, J. M. (1995). Improved model output and statistics through model consensus. *Bulletin of the American Meteorological Society*, 76(7):1157–1164.

[Voyant et al., 2012] Voyant, C., Muselli, M., Paoli, C., and Nivet, M.-L. (2012). Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. *Energy*, 39(1):341–355.

[Voyant et al., 2017] Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., and Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582.

[Wacker et al., 2015] Wacker, S., Gröbner, J., Zysset, C., Diener, L., Tzoumanikas, P., Kazantzidis, A., Vuilleumier, L., Stöckli, R., Nyeki, S., and Kämpfer, N. (2015). Cloud observations in Switzerland using hemispherical sky cameras. *Journal of Geophysical Research*, 120(2):695–707.

[Wan et al., 2014] Wan, C., Xu, Z., Pinson, P., Dong, Z. Y., and Wong, K. P. (2014). Probabilistic forecasting of wind power generation using extreme learning machine. *IEEE Transactions on Power Systems*, 29(3):1033–1044.

[Weinberger and Saul, 2009] Weinberger, K. Q. and Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10:207–244.

[Wess et al., 1994] Wess, S., Althoff, K. ., and Derwand, G. (1994). *Using k-d trees to improve the retrieval step in case-based reasoning*, volume 837 LNAI of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer.

[Wey and Schroedter-Homscheidt, 2013] Wey, E. and Schroedter-Homscheidt, M. (2013). APOLLO cloud product statistics. In *Energy Procedia*, volume 49, pages 2414–2421.

[Wiegner et al., 2014] Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W. (2014). What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET. *Atmospheric Measurement Techniques*, 7(7):1979–1997.

[Wild et al., 2013] Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., and König-Langlo, G. (2013). The global energy balance from a surface perspective. *Climate Dynamics*, 40(11-12):3107–3134.

[William et al., 2008] William, C. S., Joseph, B., Jimy, D., David, O., Dale, M., Michael, G., Huang, X., Wang, W., and Jordan, G. (2008). A description of the advanced research wrf version 3. *NCAR technical note*, 126.

[Wolff et al., 2016] Wolff, B., Lorenz, E., and Kramer, O. (2016). Statistical learning for short-term photovoltaic power predictions. In *Computational sustainability*, pages 31–45. Springer.

[World Meteorological Organization , 2012] World Meteorological Organization (2012). World Meteorological Organization/World Weather Research Programme (WMO/WWRP). *Recommended methods for evaluating cloud and related parameters World Weather Research Programme (WWRP)/Working Group on Numerical Experimentation (WGNE) Joint Working Group on Forecast Verification Research (JWGFVR)*.

[Xiao et al., 2015] Xiao, L., Wang, J., Dong, Y., and Wu, J. (2015). Combined forecasting models for wind energy forecasting: A case study in china. *Renewable and Sustainable Energy Reviews*, 44:271–288.

[Xie and Arkin, 1996] Xie, P. and Arkin, P. A. (1996). Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *Journal of Climate*, 9(4):840–858.

[Yadav and Chandel, 2014] Yadav, A. K. and Chandel, S. S. (2014). Solar radiation prediction using artificial neural network techniques: A review. *Renewable and Sustainable Energy Reviews*, 33:772–781.

[Ye et al., 2017] Ye, L., Cao, Z., and Xiao, Y. (2017). DeepCloud: Ground-based cloud image categorization using deep convolutional features. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10):5729–5740.

[Zamo et al., 2014a] Zamo, M., Mestre, O., Arbogast, P., and Pannekoucke, O. (2014a). A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part i: Deterministic forecast of hourly production. *Solar Energy*, 105:792–803.

[Zamo et al., 2014b] Zamo, M., Mestre, O., Arbogast, P., and Pannekoucke, O. (2014b). A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. part II: Probabilistic forecast of daily production. *Solar Energy*, 105:804–816.

[Zendehboudi et al., 2018] Zendehboudi, A., Baseer, M. A., and Saidur, R. (2018). Application of support vector machine models for forecasting solar and wind energy resources: A review. *Journal of Cleaner Production*, 199:272–285.

[Zhang et al., 2017] Zhang, J., Zhao, L., Deng, S., Xu, W., and Zhang, Y. (2017). A critical review of the models used to estimate solar radiation. *Renewable and Sustainable Energy Reviews*, 70:314–329.

[Zhen et al., 2015] Zhen, Z., Wang, F., Sun, Y., Mi, Z., Liu, C., Wang, B., and Lu, J. (2015). SVM based cloud classification model using total sky images for PV power forecasting. In *2015 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2015*.